

Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11) **EP 0 841 394 A2**

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:  
13.05.1998 Bulletin 1998/20

(51) Int Cl.<sup>6</sup>: C12N 15/31, C07K 14/31

(21) Application number: 97307485.9

(22) Date of filing: 24.09.1997

(84) Designated Contracting States:  
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC  
NL PT SE

(30) Priority: 24.09.1996 US 27032 P

(71) Applicants:  
• SMITHKLINE BEECHAM CORPORATION  
Philadelphia Pennsylvania 19103 (US)  
• SMITHKLINE BEECHAM PLC  
Brentford, Middlesex TW8 9EP (GB)

(72) Inventors:  
• Black, Michael T., SmithKline Beecham Pharma.  
Collegeville, Pennsylvania 19426-0989 (US)  
• Hodgson, John Edward,  
SmithKline Beecham Pharma.  
Collegeville, Pennsylvania 19426-0989 (US)  
• Knowles, David J. C.,  
SmithKline Beecham Pharma.  
Harlow, Essex CM19 5AW (GB)  
• Reichard, Raymond W.,  
SmithKline Beecham Pharma.  
King of Prussia, Pennsylvania 19406 (US)

- Nicholas, Richard O.,  
SmithKline Beecham Pharma.  
Collegeville, Pennsylvania 19426-0989 (US)
- Burnham, Martin K. R.,  
SmithKline Beecham Pharma.  
Collegeville, Pennsylvania 19426-0989 (US)
- Pratt, Julie M., SmithKline Beecham Pharma.  
Harlow, Essex CM19 5AW (GB)
- Rosenberg, Martin,  
SmithKline Beecham Pharma.  
King of Prussia, Pennsylvania 19406 (US)
- Ward, Judith M., SmithKline Beecham Pharma.  
Harlow, Essex CM19 5AW (GB)
- Lonetto, Michael A.,  
SmithKline Beecham Pharma.  
Collegeville, Pennsylvania 19426-0989 (US)

(74) Representative: Giddings, Peter John, Dr. et al  
SmithKline Beecham plc  
Corporate Intellectual Property,  
Two New Horizons Court  
Brentford, Middlesex TW8 9EP (GB)

(54) **Staphylococcus aureus polynucleotides, polypeptides and their uses**

(57) The invention provides novel polypeptides and polynucleotides encoding such polypeptides and meth-

ods for producing such polypeptides by recombinant techniques. Also provided are methods for utilizing such polypeptides to screen for antibacterial compounds.

EP 0 841 394 A2

# Description

## RELATED APPLICATIONS

This application claims benefit of U.S. Provisional Application Number 60/027,032, filed September 24, 1996.

## FIELD OF THE INVENTION

This invention relates to newly identified polynucleotides and polypeptides, and their production and uses, as well as their variants, agonists and antagonists, and their uses. In particular, in these and in other regards, the invention relates to novel polynucleotides and polypeptides set forth in the Sequence Listing.

## BACKGROUND OF THE INVENTION

The Staphylococci make up a medically important genera of microbes. They are known to produce two types of disease, invasive and toxigenic. Invasive infections are characterized generally by abscess formation effecting both skin surfaces and deep tissues. *Staphylococcus aureus* (herein "*S. aureus*") is the second leading cause of bacteremia in cancer patients. Osteomyelitis, septic arthritis, septic thrombophlebitis and acute bacterial endocarditis are also relatively common. There are at least three clinical conditions resulting from the toxigenic properties of Staphylococci. The manifestation of these diseases result from the actions of exotoxins as opposed to tissue invasion and bacteremia. These conditions include: Staphylococcal food poisoning, scalded skin syndrome and toxic shock syndrome.

While certain Staphylococcal proteins associated with pathogenicity have been identified, e.g., coagulase, hemolysins, leucocidins and exo and enterotoxins, very little is known concerning the temporal expression of such genes during infection and disease progression in a mammalian host. Discovering the sets of genes the bacterium is likely to be expressing at the different stages of infection, particularly when an infection is established, provides critical information for the screening and characterization of novel antibacterials which can interrupt pathogenesis. In addition to providing a fuller understanding of known proteins, such an approach will identify previously unrecognised targets.

GUG is used as an initiating nucleotide, rather than ATG, for a significant number of mRNA's in both Gram positive and Gram negative bacteria. Statistics on the frequency of NTG codons in the start codon for several bacterial species are available on line via computer ([http://lbcchem.otago.ac.nz:800/Transterm/home\\_page.html](http://lbcchem.otago.ac.nz:800/Transterm/home_page.html)).

A discussion of initiation codons in *B. subtilis* is set forth in Vellanoweth, RL.1993 in *Bacillus subtilis and other Gram Positive Bacteria, Biochemistry, Physiology and Molecular Genetics*, Sonenshein, Hoch, Losick Eds. Amer. Soc. Microbiol, Washington DC. p. 699-711. Vellanoweth indicates a major difference between *B. subtilis* and the gram-negative organisms is in the choice of initiation codon. 91% of the sequenced *E. coli* genes start with AUG. By contrast, about 30% of *B. subtilis* and other clostridial branch genes start with UUG or GUG. Moreover, CUG functions as a start codon in *B. subtilis*. Mutations of an AUG initiation codon to GUG or UUG often cause decreased expression in *B. subtilis* and *E. coli*. Generally, translation efficiency is higher with AUG initiation codons. A strong Shine-Delgarno ribosome binding site, however, can compensate almost fully for a weak initiation codon. It has been reported that genes with a range of expression levels have initiation codons other than ATG in gram positives (Vellanoweth, RL.1993 in *Bacillus subtilis and other Gram Positive Bacteria, Biochemistry, Physiology and Molecular Genetics*, Sonenshein, Hoch, Losick Eds. Amer. Soc. Microbiol, Washington DC. p. 699-711).

Provided herein are ORF sequences from genes possessing GUG initiation codons and proteins expressed therefrom to be used for screening for antimicrobial compounds. Clearly, there is a need for polypeptide and polynucleotide sequences that may be used to screen for antimicrobial compound and which may also be used to determine the roles of such sequences in pathogenesis of infection, dysfunction and disease. There is also need, therefore, for identification and characterization of such sequences which may play a role in preventing, ameliorating or correcting infections, dysfunctions or diseases.

The polypeptides of the invention have amino acid sequence homology to a known protein(s) as set forth in Table 1.

## SUMMARY OF THE INVENTION

It is an object of the invention to provide polypeptides that have been identified as novel polypeptides by homology between an amino acid sequence selected from the group consisting of the sequences set out in the Sequence Listing and a known amino acid sequence or sequences of other proteins such as the protein identities listed in Table 1.

It is a further object of the invention to provide polynucleotides that encode novel polypeptides, particularly polynucleotides that encode polypeptides of *Staphylococcus aureus*.

In a particularly preferred embodiment of the invention the polynucleotide comprises a region encoding a polypeptide comprising a sequence selected from the group consisting of the sequences set out in the Sequence

Listing, or a variant of any of these sequences.

In another particularly preferred embodiment of the invention there is a novel protein from *Staphylococcus aureus* comprising an amino acid sequence selected from the group consisting of the sequences set out in the Sequence Listing, or a variant of any of these sequences.

In accordance with another aspect of the invention there is provided an isolated nucleic acid molecule encoding a mature polypeptide expressible by the *Staphylococcus aureus* WCUH29 strain contained in the deposited strain.

A further aspect of the invention there are provided isolated nucleic acid molecules encoding a polypeptide of the invention, particularly *Staphylococcus aureus* polypeptide, and including mRNAs, cDNAs, genomic DNAs. Further embodiments of the invention include biologically, diagnostically, prophylactically, clinically or therapeutically useful variants thereof, and compositions comprising the same.

In accordance with another aspect of the invention, there is provided the use of a polynucleotide of the invention for therapeutic or prophylactic purposes, in particular genetic immunization. Among the particularly preferred embodiments of the invention are naturally occurring allelic variants of a polypeptide of the invention and polypeptides encoded thereby.

Another aspect of the invention there are provided novel polypeptides of *Staphylococcus aureus* as well as biologically, diagnostically, prophylactically, clinically or therapeutically useful variants thereof, and compositions comprising the same.

Among the particularly preferred embodiments of the invention are variants of the polypeptides of the invention encoded by naturally occurring alleles of their genes.

In a preferred embodiment of the invention there are provided methods for producing the aforementioned polypeptides.

In accordance with yet another aspect of the invention, there are provided inhibitors to such polypeptides, useful as antibacterial agents, including, for example, antibodies.

In accordance with certain preferred embodiments of the invention, there are provided products, compositions and methods for assessing expression of the polypeptides and polynucleotides of the invention, treating disease, for example, infections of the upper respiratory tract (e.g., otitis media, bacterial tracheitis, acute epiglottitis, thyroiditis), lower respiratory (e.g., empyema, lung abscess), cardiac (e.g., infective endocarditis), gastrointestinal (e.g., secretory diarrhoea, splenic abscess, retroperitoneal abscess), CNS (e.g., cerebral abscess), eye (e.g., blepharitis, conjunctivitis, keratitis, endophthalmitis, preseptal and orbital cellulitis, dacryocystitis), kidney and urinary tract (e.g., epididymitis, intrarenal and perinephric abscess, toxic shock syndrome), skin (e.g., impetigo, folliculitis, cutaneous abscesses, cellulitis, wound infection, bacterial myositis) bone and joint (e.g., septic arthritis, osteomyelitis), assaying genetic variation, and administering a polypeptide or polynucleotide of the invention to an organism to raise an immunological response against a bacteria, especially a *Staphylococcus aureus* bacteria.

In accordance with certain preferred embodiments of this and other aspects of the invention there are provided polynucleotides that hybridize to a polynucleotide sequence of the invention, particularly under stringent conditions.

In certain preferred embodiments of the invention there are provided antibodies against polypeptides of the invention.

In other embodiments of the invention there are provided methods for identifying compounds which bind to or otherwise interact with and inhibit or activate an activity of a polypeptide or polynucleotide of the invention comprising: contacting a polypeptide or polynucleotide of the invention with a compound to be screened under conditions to permit binding to or other interaction between the compound and the polypeptide or polynucleotide to assess the binding to or other interaction with the compound, such binding or interaction being associated with a second component capable of providing a detectable signal in response to the binding or interaction of the polypeptide or polynucleotide with the compound; and determining whether the compound binds to or otherwise interacts with and activates or inhibits an activity of the polypeptide or polynucleotide by detecting the presence or absence of a signal generated from the binding or interaction of the compound with the polypeptide or polynucleotide.

In accordance with yet another aspect of the invention, there are provided agonists and antagonists of the polypeptides and polynucleotides of the invention, preferably bacteriostatic or bacteriocidal agonists and antagonists.

In a further aspect of the invention there are provided compositions comprising a polynucleotide or a polypeptide of the invention for administration to a cell or to a multicellular organism.

Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following descriptions and from reading the other parts of the present disclosure.

## GLOSSARY

The following definitions are provided to facilitate understanding of certain terms used frequently herein.

"Host cell" is a cell which has been transformed or transfected, or is capable of transformation or transfection by

an exogenous polynucleotide sequence.

"Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" and "similarity" can be readily calculated by known methods, including but not limited to those described in (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., *SIAM J. Applied Math.*, 48: 1073 (1988). Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, the GCG program package (Devereux, J., et al., *Nucleic Acids Research* 12(1): 387 (1984)), BLASTP, BLASTN, and FASTA (Altschul, S.F. et al., *J. Molec. Biol.* 215: 403-410 (1990)). The BLAST X program is publicly available from NCBI and other sources (*BLAST Manual*, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., *J. Mol. Biol.* 215: 403-410 (1990)). As an illustration, by a polynucleotide having a nucleotide sequence having at least, for example, 95% "identity" to a reference nucleotide sequence it is intended that the nucleotide sequence of the tested polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. These mutations of the reference sequence may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence. Analogously, by a polypeptide having an amino acid sequence having at least, for example, 95% identity to a reference amino acid sequence is intended that the test amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the reference amino acid. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

"Isolated" means altered "by the hand of man" from its natural state, i.e., if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living organism is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

"Polynucleotide(s)" generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotide(s)" include, without limitation, single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions or single-, double- and triple-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded, or triple-stranded regions, or a mixture of single- and double-stranded regions. In addition, "polynucleotide" as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide. As used herein, the term "polynucleotide(s)" also includes DNAs or RNAs as described above that contain one or more modified bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "polynucleotide(s)" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein. It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term "polynucleotide(s)" as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including, for example, simple and complex cells. "Polynucleotide(s)" also embraces short polynucleotides often referred to as oligonucleotide(s).



"Polypeptide(s)" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds. "Polypeptide(s)" refers to both short chains, commonly referred to as peptides, oligopeptides and oligomers and to longer chains generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene encoded amino acids. "Polypeptide(s)" include those modified either by natural processes, such as processing and other post-translational modifications, but also by chemical modification techniques. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature, and they are well known to those of skill in the art. It will be appreciated that the same type of modification may be present in the same or varying degree at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains, and the amino or carboxyl termini. Modifications include, for example, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins, such as arginylation, and ubiquitination. See, for instance, *PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES*, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993) and Wold, F., *Posttranslational Protein Modifications: Perspectives and Prospects*, pgs. 1-12 in *POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS*, B. C. Johnson, Ed., Academic Press, New York (1983); Seifter et al., *Meth. Enzymol.* 182:626-646 (1990) and Rattan et al., *Protein Synthesis: Posttranslational Modifications and Aging*, Ann. N.Y. Acad. Sci. 663: 48-62 (1992). Polypeptides may be branched or cyclic, with or without branching. Cyclic, branched and branched circular polypeptides may result from post-translational natural processes and may be made by entirely synthetic methods, as well.

"Variant(s)" as the term is used herein, is a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide respectively, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques, by direct synthesis, and by other recombinant methods known to skilled artisans.

#### DESCRIPTION OF THE INVENTION

Each of polynucleotide and polypeptide sequences provided herein may be used in the discovery and development of antibacterial compounds. Upon expression of the sequences with the appropriate initiation and termination codons the encoded polypeptide can be used as a target for the screening of antimicrobial drugs. Additionally, the DNA sequences encoding preferably the amino terminal regions of the encoded protein or the Shine-Delgarno region can be used to construct antisense sequences to control the expression of the coding sequence of interest. Furthermore, many of the sequences disclosed herein also provide regions upstream and downstream from the encoding sequence. These sequences are useful as a source of regulatory elements for the control of bacterial gene expression. Such sequences are conveniently isolated by restriction enzyme action or synthesized chemically and introduced, for example, into promoter identification strains. These strains contain a reporter structural gene sequence located downstream from a restriction site such that if an active promoter is inserted, the reporter gene will be expressed.

Although each of the sequences may be employed as described above, this invention also provides several means for identifying particularly useful target genes. The first of these approaches entails searching appropriate databases for sequence matches in related organisms. Thus, if a homologue exists, the Streptococcal-like form of this gene would likely play an analogous role. For example, a Streptococcal protein identified as homologous to a cell surface protein in another organism would be useful as a vaccine candidate. To the extent such homologies have been identified for the sequences disclosed herein they are reported along with the encoding sequence.

Each of the DNA sequences provided herein may be used in the discovery and development of antibacterial com-

pounds. Because each of the sequences contains an open reading frame (ORF) with an appropriate initiation and termination codons, the encoded protein upon expression can be used as a target for the screening of antimicrobial drugs. Additionally, the DNA sequences encoding the amino terminal regions of the encoded protein can be used to construct antisense sequences to control the expression of the coding sequence of interest. Furthermore, many of the sequences disclosed herein also provide regions upstream and downstream from the encoding sequence. These sequences are useful as a source of regulatory elements for the control of bacterial gene expression. Such sequences are conveniently isolated by restriction enzyme action or synthesized chemically and introduced, for example, into promoter identification strains. These strains contain a reporter structural gene sequence located downstream from a restriction site such that if an active promoter is inserted, the reporter gene will be expressed.

It is believed that bacteria possess a number of ways of regulating gene expression levels, especially in subtle degrees, and the interplay between ribosome binding site and initiation codon is utilized for this purpose for these genes. It is also believed that such genes will be important targets for antimicrobial drug discovery, particularly since pathogenesis genes are believed undergo gene expression regulation during in the pathogenesis process. Therefore, the invention provides ORF sequences possessing a GTG (GUG) initiation codon and protein targets expressed therefrom.

Although each of the sequences may be employed as described above, this invention also provides several means for identifying particularly useful target genes. The first of these approaches entails searching appropriate databases for sequence matches in related organisms. Thus, if a homologue exists, the Staphylococcal-like form of this gene would likely play an analogous role. For example, a Staphylococcal protein identified as homologous to a cell surface protein in another organism would be useful as a vaccine candidate. To the extent such homologies have been identified for the sequences disclosed herein they are reported along with the encoding sequence.

#### ORF Gene Expression

Recently techniques have become available to evaluate temporal gene expression in bacteria, particularly as it applies to viability under laboratory and infection conditions. A number of methods can be used to identify genes which are essential to survival *per se*, or essential to the establishment/maintenance of an infection. Identification of an ORF unknown by one of these methods yields additional information about its function and permits the selection of such an ORF for further development as a screening target. Briefly, these approaches include:

**1) Signature Tagged Mutagenesis (STM):** This technique is described by Hensel *et al.*, *Science* 269: 400-403 (1995), the contents of which is incorporated by reference for background purposes. Signature tagged mutagenesis identifies genes necessary for the establishment/maintenance of infection in a given infection model.

The basis of the technique is the random mutagenesis of target organism by various means (e.g., transposons) such that unique DNA sequence tags are inserted in close proximity to the site of mutation. The tags from a mixed population of bacterial mutants and bacteria recovered from an infected hosts are detected by amplification, radiolabeling and hybridisation analysis. Mutants attenuated in virulence are revealed by absence of the tag from the pool of bacteria recovered from infected hosts.

In *Staphylococcus aureus*, because the transposon system is less well developed, a more efficient way of creating the tagged mutants is to use the insertion-duplication mutagenesis technique as described by Morrison *et al.*, *J. Bacteriol.* 159:870 (1984) the contents of which is incorporated by reference for background purposes.

**2) In Vivo Expression Technology (IVET):** This technique is described by Camilli *et al.*, *Proc. Nat'l. Acad. Sci. USA.* 91:2634-2638 (1994), the contents of which is incorporated by reference for background purposes. IVET identifies genes up-regulated during infection when compared to laboratory cultivation, implying an important role in infection. ORF identified by this technique are implied to have a significant role in infection establishment/maintenance.

In this technique random chromosomal fragments of target organism are cloned upstream of a promoter-less recombinase gene in a plasmid vector. This construct is introduced into the target organism which carries an antibiotic resistance gene flanked by resolvase sites. Growth in the presence of the antibiotic removes from the population those fragments cloned into the plasmid vector capable of supporting transcription of the recombinase gene and therefore have caused loss of antibiotic resistance. The resistant pool is introduced into a host and at various times after infection bacteria may be recovered and assessed for the presence of antibiotic resistance. The chromosomal fragment carried by each antibiotic sensitive bacterium should carry a promoter or portion of a gene normally upregulated during infection. Sequencing upstream of the recombinase gene allows identification of the up regulated gene.

**3) Differential display:** This technique is described by Chuang *et al.*, *J. Bacteriol.* 175:2026-2036 (1993), the contents of which is incorporated by reference for background purposes. This method identifies those genes which are expressed in an organism by identifying mRNA present using randomly-primed RT-PCR. By comparing pre-

infection and post infection profiles, genes up and down regulated during infection can be identified and the RT-PCR product sequenced and matched to ORF 'unknowns'.

**4) Generation of conditional lethal mutants by transposon mutagenesis:** This technique, described by de Lorenzo, V. *et al.*, *Gene* 123:17-24 (1993); Neuwald, A. F. *et al.*, *Gene* 125: 69-73(1993); and Takiff, H. E. *et al.*, *J. Bacteriol.* 174:1544-1553(1992), the contents of which is incorporated by reference for background purposes, identifies genes whose expression are essential for cell viability.

In this technique transposons carrying controllable promoters, which provide transcription outward from the transposon in one or both directions, are generated. Random insertion of these transposons into target organisms and subsequent isolation of insertion mutants in the presence of inducer of promoter activity ensures that insertions which separate promoter from coding region of a gene whose expression is essential for cell viability will be recovered. Subsequent replica plating in the absence of inducer identifies such insertions, since they fail to survive. Sequencing of the flanking regions of the transposon allows identification of site of insertion and identification of the gene disrupted. Close monitoring of the changes in cellular processes/morphology during growth in the absence of inducer yields information on likely function of the gene. Such monitoring could include flow cytometry (cell division, lysis, redox potential, DNA replication), incorporation of radiochemically labeled precursors into DNA, RNA, protein, lipid, peptidoglycan, monitoring reporter enzyme gene fusions which respond to known cellular stresses.

**5) Generation of conditional lethal mutants by chemical mutagenesis:** This technique is described by Beckwith, J., *Methods in Enzymology* 204:

3-18(1991), the contents of which are incorporated herein by reference for background purposes. In this technique random chemical mutagenesis of target organism, growth at temperature other than physiological temperature (permissive temperature) and subsequent replica plating and growth at different temperature (e.g. 42°C to identify ts, 25°C to identify cs) are used to identify those isolates which now fail to grow (conditional mutants). As above close monitoring of the changes upon growth at the non-permissive temperature yields information on the function of the mutated gene. Complementation of conditional lethal mutation by library from target organism and sequencing of complementing gene allows matching with unknown ORF.

**6) RT-PCR:** *Staphylococcus aureus* messenger RNA is isolated from bacterial infected tissue e.g. 48 hour murine lung infections, and the amount of each mRNA species assessed by reverse transcription of the RNA sample primed with random hexanucleotides followed by PCR with gene specific primer pairs. The determination of the presence and amount of a particular mRNA species by quantification of the resultant PCR product provides information on the bacterial genes which are transcribed in the infected tissue. Analysis of gene transcription can be carried out at different times of infection to gain a detailed knowledge of gene regulation in bacterial pathogenesis allowing for a clearer understanding of which gene products represent targets for screens for novel antibacterials. Because of the gene specific nature of the PCR primers employed it should be understood that the bacterial mRNA preparation need not be free of mammalian RNA. This allows the investigator to carry out a simple and quick RNA preparation from infected tissue to obtain bacterial mRNA species which are very short lived in the bacterium (in the order of 2 minute half-lives). Optimally the bacterial mRNA is prepared from infected murine lung tissue by mechanical disruption in the presence of TRIzol (GIBCO-BRL) for very short periods of time, subsequent processing according to the manufacturers of TRIzol reagent and DNAase treatment to remove contaminating DNA. Preferably the process is optimised by finding those conditions which give a maximum amount of *Staphylococcus aureus* 16S ribosomal RNA as detected by probing Northern blots with a suitably labelled sequence specific oligonucleotide probe. Typically a 5' dye labelled primer is used in each PCR primer pair in a PCR reaction which is terminated optimally between 8 and 25 cycles. The PCR products are separated on 6% polyacrylamide gels with detection and quantification using GeneScanner (manufactured by ABI).

Each of these techniques may have advantages or disadvantage depending on the particular application. The skilled artisan would choose the approach that is the most relevant with the particular end use in mind.

Use of the of these technologies when applied to the ORFs of the present invention enables identification of bacterial proteins expressed during infection, inhibitors of which would have utility in anti-bacterial therapy.

The invention relates to novel polypeptides and polynucleotides as described in greater detail below. In particular, the invention relates to polypeptides and polynucleotides of *Staphylococcus aureus*, which is related by amino acid sequence homology to known polypeptide as set forth in Table 1. The invention relates especially to compounds having the nucleotide and amino acid sequence selected from the group consisting of the sequences set out in the Sequence Listing, and to the nucleotide sequences of the DNA in the deposited strain and amino acid sequences encoded thereby.

#### Deposited materials

*S. aureus* WCUH 29 has been deposited at the National Collection of Industrial and Marine Bacteria Ltd. (NCIMB),

Aberdeen, Scotland under number NCIMB 40771 on 11 September 1995.

The *Staphylococcus aureus* strain deposit is referred to herein as "the deposited strain" or as "the DNA of the deposited strain."

The deposited strain contains the full length genes comprising the polynucleotides set forth in the Sequence Listing. The sequence of the polynucleotides contained in the deposited strain, as well as the amino acid sequence of the polypeptide encoded thereby, are controlling in the event of any conflict with any description of sequences herein.

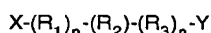
The deposit of the deposited strain has been made under the terms of the Budapest Treaty on the International Recognition of the Deposit of Micro-organisms for Purposes of Patent Procedure. The strain will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. The deposited strain is provided merely as convenience to those of skill in the art and is not an admission that a deposit is required for enablement, such as that required under 35 U.S.C. §112.

A license may be required to make, use or sell the deposited strain, and compounds derived therefrom, and no such license is hereby granted.

### Polypeptides

The polypeptides of the invention include the polypeptides set forth in the Sequence Listing (in particular the mature polypeptide) as well as polypeptides and fragments, particularly those which have the biological activity of a polypeptide of the invention, and also those which have at least 50%, 60% or 70% identity to a polypeptide sequence selected from the group consisting of the sequences set out in the Sequence Listing or the relevant portion, preferably at least 80% identity to a polypeptide sequence selected from the group consisting of the sequences set out in the Sequence Listing, and more preferably at least 90% similarity (more preferably at least 90% identity) to a polypeptide sequence selected from the group consisting of the sequences set out in the Sequence Listing, and still more preferably at least 95% similarity (still more preferably at least 95% identity) to a polypeptide sequence selected from the group consisting of the sequences set out in the Sequence Listing, and also include portions of such polypeptides with such portion of the polypeptide generally containing at least 30 amino acids and more preferably at least 50 amino acids.

The invention also includes polypeptides of the formula:



wherein, at the amino terminus, X is hydrogen, and at the carboxyl terminus, Y is hydrogen or a metal,  $R_1$  and  $R_3$  are any amino acid residue, n is an integer between 1 and 2000, and  $R_2$  is an amino acid sequence of the invention, particularly an amino acid sequence selected from the group set forth in the Sequence Listing. In the formula above  $R_2$  is oriented so that its amino terminal residue is at the left, bound to  $R_1$ , and its carboxy terminal residue is at the right, bound to  $R_3$ . Any stretch of amino acid residues denoted by either R group, where R is greater than 1, may be either a heteropolymer or a homopolymer, preferably a heteropolymer. In preferred embodiments n is an integer between 1 and 1000 or 2000.

A fragment is a variant polypeptide having an amino acid sequence that entirely is the same as part but not all of the amino acid sequence of the aforementioned polypeptides. As with polypeptides, fragments may be "free-standing," or comprised within a larger polypeptide of which they form a part or region, most preferably as a single continuous region, a single larger polypeptide.

Preferred fragments include, for example, truncation polypeptides having a portion of the amino acid sequence of the Sequence Listing, or of variants thereof, such as a continuous series of residues that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus. Degradation forms of the polypeptides of the invention in a host cell, particularly a *Staphylococcus aureus*, are also preferred. Further preferred are fragments characterized by structural or functional attributes such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions.

Also preferred are biologically active fragments which are those fragments that mediate activities of polypeptides of the invention, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also included are those fragments that are antigenic or immunogenic in an animal, especially in a human. Particularly preferred are fragments comprising receptors or domains of enzymes that confer a function essential for viability of *Staphylococcus aureus* or the ability to initiate, or maintain cause disease in an individual, particularly a human.

Variants that are fragments of the polypeptides of the invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, these variants may be employed as intermediates for producing the full-length polypeptides of the invention.

In addition to the standard single and triple letter representations for amino acids, the term "X" or "Xaa" is also used. "X" and "Xaa" mean that any of the twenty naturally occurring amino acids may appear at such a designated position in the polypeptide sequence.

## 5 Polynucleotides

The nucleotide sequences disclosed herein can be obtained by synthetic chemical techniques known in the art or can be obtained from *S. aureus* WCUH 29 by probing a DNA preparation with probes constructed from the particular sequences disclosed herein. Alternatively, oligonucleotides derived from a disclosed sequence can act as PCR primers in a process of PCR-based cloning of the sequence from a bacterial genomic source. It is recognised that such sequences will also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained.

To obtain the polynucleotide encoding the protein using the DNA sequence given herein typically a library of clones of chromosomal DNA of *S. aureus* WCUH 29 in *E. coli* or some other suitable host is probed with a radiolabelled oligonucleotide, preferably a 17mer or longer, derived from the partial sequence. Clones carrying DNA identical to that of the probe can then be distinguished using high stringency washes. By sequencing the individual clones thus identified with sequencing primers designed from the original sequence it is then possible to extend the sequence in both directions to determine the full gene sequence. Conveniently such sequencing is performed using denatured double stranded DNA prepared from a plasmid clone. Suitable techniques are described by Maniatis, T., Fritsch, E.F. and Sambrook, J. in *MOLECULAR CLONING, A Laboratory Manual*, 2nd edition, 1989, Cold Spring Harbor Laboratory (see: Screening By Hybridization 1.90 and Sequencing Denatured Double-Stranded DNA Templates 13.70).

Moreover, another aspect of the invention relates to isolated polynucleotides that encode the polypeptides of the invention having a deduced amino acid sequence selected from the group consisting of the sequences in the Sequence Listing and polynucleotides closely related thereto and variants thereof.

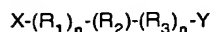
Using the information provided herein, such as the polynucleotide sequences set out in the Sequence Listing, a polynucleotide of the invention encoding polypeptide may be obtained using standard cloning and screening methods, such as those for cloning and sequencing chromosomal DNA fragments from bacteria using *Staphylococcus aureus* WCUH29 cells as starting material, followed by obtaining a full length clone. For example, to obtain a polynucleotide sequence of the invention, such as a sequence set forth in the Sequence Listing, typically a library of clones of chromosomal DNA of *Staphylococcus aureus* WCUH29 in *E. coli* or some other suitable host is probed with a radiolabeled oligonucleotide, preferably a 17-mer or longer, derived from a partial sequence. Clones carrying DNA identical to that of the probe can then be distinguished using stringent conditions. By sequencing the individual clones thus identified with sequencing primers designed from the original sequence it is then possible to extend the sequence in both directions to determine the full gene sequence. Conveniently, such sequencing is performed using denatured double stranded DNA prepared from a plasmid clone. Suitable techniques are described by Maniatis, T., Fritsch, E.F. and Sambrook et al., *MOLECULAR CLONING, A LABORATORY MANUAL*, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989). (see in particular Screening By Hybridization 1.90 and Sequencing Denatured Double-Stranded DNA Templates 13.70). Illustrative of the invention, the polynucleotides set out in the Sequence Listing were discovered in a DNA library derived from *Staphylococcus aureus* WCUH29.

The DNA sequences set out in the Sequence Listing each contains at least one open reading frame encoding a protein having at least about the number of amino acid residues set forth in the Sequence Listing. The start and stop codons of each open reading frame (herein "ORF") DNA are the first three and the last three nucleotides of each polynucleotide set forth in the Sequence Listing.

Certain polynucleotides and polypeptides of the invention are structurally related to known proteins as set forth in Table 1. These proteins exhibit greatest homology to the homologue listed in Table 1 from among the known proteins.

The invention provides a polynucleotide sequence identical over its entire length to each coding sequence in the Sequence Listing. Also provided by the invention is the coding sequence for the mature polypeptide or a fragment thereof, by itself as well as the coding sequence for the mature polypeptide or a fragment in reading frame with other coding sequence, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence. The polynucleotide may also contain non-coding sequences, including for example, but not limited to non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences, termination signals, ribosome binding sites, sequences that stabilize mRNA, introns, polyadenylation signals, and additional coding sequence which encode additional amino acids. For example, a marker sequence that facilitates purification of the fused polypeptide can be encoded. In certain embodiments of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz et al., *Proc. Natl. Acad. Sci., USA* 86: 821-824 (1989), or an HA tag (Wilson et al., *Cell* 37: 767 (1984). Polynucleotides of the invention also include, but are not limited to, polynucleotides comprising a structural gene and its naturally associated sequences that control gene expression.

The invention also includes polynucleotides of the formula:



wherein, at the 5' end of the molecule, X is hydrogen, and at the 3' end of the molecule, Y is hydrogen or a metal,  $R_1$  and  $R_3$  is any nucleic acid residue, n is an integer between 1 and 3000, and  $R_2$  is a nucleic acid sequence of the invention, particularly a nucleic acid sequence selected from the group set forth in the Sequence Listing. In the polynucleotide formula above  $R_2$  is oriented so that its 5' end residue is at the left, bound to  $R_1$ , and its 3' end residue is at the right, bound to  $R_3$ . Any stretch of nucleic acid residues denoted by either R group, where R is greater than 1, may be either a heteropolymer or a homopolymer, preferably a heteropolymer. In a preferred embodiment n is an integer between 1 and 1000, or 2000 or 3000.

The term "polynucleotide encoding a polypeptide" as used herein encompasses polynucleotides that include a sequence encoding a polypeptide of the invention, particularly a bacterial polypeptide and more particularly a polypeptide of the *Staphylococcus aureus* having an amino acid sequence set out in the Sequence Listing. The term also encompasses polynucleotides that include a single continuous region or discontinuous regions encoding the polypeptide (for example, interrupted by integrated phage or an insertion sequence or editing) together with additional regions, that also may contain coding and/or non-coding sequences.

The invention further relates to variants of the polynucleotides described herein that encode for variants of the polypeptide having the deduced amino acid sequence of the Sequence Listing. Variants that are fragments of the polynucleotides of the invention may be used to synthesize full-length polynucleotides of the invention.

Further particularly preferred embodiments are polynucleotides encoding polypeptide variants, that have the amino acid sequence of a polypeptide of the Sequence Listing in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, deleted or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, that do not alter the properties and activities of such polynucleotide.

Further preferred embodiments of the invention are polynucleotides that are at least 50%, 60% or 70% identical over their entire length to a polynucleotide encoding a polypeptide having the amino acid sequence set out in the Sequence Listing, and polynucleotides that are complementary to such polynucleotides. Alternatively, most highly preferred are polynucleotides that comprise a region that is at least 80% identical over its entire length to a polynucleotide encoding a polypeptide of the deposited strain and polynucleotides complementary thereto. In this regard, polynucleotides at least 90% identical over their entire length to the same are particularly preferred, and among these particularly preferred polynucleotides, those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred among those with at least 95%, and among these those with at least 98% and at least 99% are particularly highly preferred, with at least 99% being the more preferred.

A preferred embodiment is an isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of: a polynucleotide having at least a 50% identity to a polynucleotide encoding a polypeptide comprising the amino acid sequence of the Sequence Listing and obtained from a prokaryotic species other than *S. aureus*; and a polynucleotide encoding a polypeptide comprising an amino acid sequence which is at least 50% identical to the amino acid sequence of the Sequence Listing and obtained from a prokaryotic species other than *S. aureus*.

Preferred embodiments are polynucleotides that encode polypeptides that retain substantially the same biological function or activity as the mature polypeptide encoded by the DNA of the Sequence Listing.

The invention further relates to polynucleotides that hybridize to the herein above-described sequences. In this regard, the invention especially relates to polynucleotides that hybridize under stringent conditions to the herein above-described polynucleotides. As herein used, the terms "stringent conditions" and "stringent hybridization conditions" mean hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences. An example of stringent hybridization conditions is overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 micrograms/ml denatured, sheared salmon sperm DNA, followed by washing the hybridization support in 0.1x SSC at about 65°C. Hybridization and wash conditions are well known and exemplified in Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), particularly Chapter 11 therein.

The invention also provides a polynucleotide consisting essentially of a polynucleotide sequence obtainable by screening an appropriate library containing the complete gene for a polynucleotide sequence set forth in the Sequence Listing under stringent hybridization conditions with a probe having the sequence of said polynucleotide sequence or a fragment thereof; and isolating said DNA sequence. Fragments useful for obtaining such a polynucleotide include, for example, probes and primers described elsewhere herein.

As discussed additionally herein regarding polynucleotide assays of the invention, for instance, polynucleotides of the invention as discussed above, may be used as a hybridization probe for RNA, cDNA and genomic DNA to isolate full-length cDNAs and genomic clones encoding a polypeptide and to isolate cDNA and genomic clones of other genes that have a high sequence similarity to a polynucleotide set forth in the Sequence Listing. Such probes generally will

comprise at least 15 bases. Preferably, such probes will have at least 30 bases and may have at least 50 bases. Particularly preferred probes will have at least 30 bases and will have 50 bases or less.

For example, the coding region of each gene that comprises or is comprised by a polynucleotide set forth in the Sequence Listing may be isolated by screening using a DNA sequence provided in the Sequence Listing to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to that of a gene of the invention is then used to screen a library of cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

The polynucleotides and polypeptides of the invention may be employed, for example, as research reagents and materials for discovery of treatments of and diagnostics for disease, particularly human disease, as further discussed herein relating to polynucleotide assays.

Polynucleotides of the invention that are oligonucleotides derived from the a polynucleotide or polypeptide sequence set forth in the Sequence Listing may be used in the processes herein as described, but preferably for PCR, to determine whether or not the polynucleotides identified herein in whole or in part are transcribed in bacteria in infected tissue. It is recognized that such sequences will also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained.

The invention also provides polynucleotides that may encode a polypeptide that is the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, may allow protein transport, may lengthen or shorten protein half-life or may facilitate manipulation of a protein for assay or production, among other things. As generally is the case *in vivo*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

A precursor protein, having the mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are removed such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

In addition to the standard A, G, C, T/U representations for nucleic acid bases, the term "N" is also used. "N" means that any of the four DNA or RNA bases may appear at such a designated position in the DNA or RNA sequence, except it is preferred that N is not a base that when taken in combination with adjacent nucleotide positions, when read in the correct reading frame, would have the effect of generating a premature termination codon in such reading frame.

In sum, a polynucleotide of the invention may encode a mature protein, a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences that are not the leader sequences of a preprotein, or a preproprotein, which is a precursor to a proprotein, having a leader sequence and one or more prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

### Vectors, host cells, expression

The invention also relates to vectors that comprise a polynucleotide or polynucleotides of the invention, host cells that are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the invention.

For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof or polynucleotides of the invention. Introduction of a polynucleotide into the host cell can be effected by methods described in many standard laboratory manuals, such as Davis et al., *BASIC METHODS IN MOLECULAR BIOLOGY*, (1986) and Sambrook et al., *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), such as, calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction and infection.

Representative examples of appropriate hosts include bacterial cells, such as staphylococci, staphylococci, enterococci *E. coli*, streptomyces and *Bacillus subtilis* cells; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293 and Bowes melanoma cells; and plant cells.

A great variety of expression systems can be used to produce the polypeptides of the invention. Such vectors include, among others, chromosomal, episomal and virus-derived vectors, *e.g.*, vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression system constructs may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to

maintain, propagate or express polynucleotides and/or to express a polypeptide in a host may be used for expression in this regard. The appropriate DNA sequence may be inserted into the expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL*, (*supra*).

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

Polypeptides of the invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography, and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding protein may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

#### Diagnostic Assays

This invention is also related to the use of the polynucleotides of the invention for use as diagnostic reagents. Detection of such polynucleotides in a eukaryote, particularly a mammal, and especially a human, will provide a diagnostic method for diagnosis of a disease. Eukaryotes (herein also "individual(s)"), particularly mammals, and especially humans, infected with an organism comprising a gene of the invention may be detected at the nucleic acid level by a variety of techniques.

Nucleic acids for diagnosis may be obtained from an infected individual's cells and tissues, such as bone, blood, muscle, cartilage, and skin. Genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR or other amplification technique prior to analysis. RNA or cDNA may also be used in the same ways. Using amplification, characterization of the species and strain of prokaryote present in an individual, may be made by an analysis of the genotype of the prokaryote gene. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the genotype of a reference sequence. Point mutations can be identified by hybridizing amplified DNA to labeled polynucleotide sequences of the invention. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by alterations in the electrophoretic mobility of the DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing. See, *e.g.*, Myers *et al.*, *Science*, 230: 1242 (1985). Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase and S1 protection or a chemical cleavage method. See, *e.g.*, Cotton *et al.*, *Proc. Natl. Acad. Sci., USA*, 85: 4397-4401 (1985).

Cells carrying mutations or polymorphisms in the gene of the invention may also be detected at the DNA level by a variety of techniques, to allow for serotyping, for example. For example, RT-PCR can be used to detect mutations. It is particularly preferred to use RT-PCR in conjunction with automated detection systems, such as, for example, GeneScan. RNA or cDNA may also be used for the same purpose, PCR or RT-PCR. As an example, PCR primers complementary to a nucleic acid encoding a polypeptide of the invention can be used to identify and analyze mutations. These primers may be used for, among other things, amplifying a DNA of the invention isolated from a sample derived from an individual. The primers may be used to amplify the gene isolated from an infected individual such that the gene may then be subject to various techniques for elucidation of the DNA sequence. In this way, mutations in the DNA sequence may be detected and used to diagnose infection and to serotype and/or classify the infectious agent.

The invention further provides a process for diagnosing disease, preferably bacterial infections, more preferably infections by *Staphylococcus aureus*, and most preferably disease, such as, infections of the upper respiratory tract (*e.g.*, otitis media, bacterial tracheitis, acute epiglottitis, thyroiditis), lower respiratory (*e.g.*, empyema, lung abscess), cardiac (*e.g.*, infective endocarditis), gastrointestinal (*e.g.*, secretory diarrhoea, splenic abscess, retroperitoneal abscess), CNS (*e.g.*, cerebral abscess), eye (*e.g.*, blepharitis, conjunctivitis, keratitis, endophthalmitis, preseptal and orbital cellulitis, dacryocystitis), kidney and urinary tract (*e.g.*, epididymitis, intrarenal and perinephric abscess, toxic shock syndrome), skin (*e.g.*, impetigo, folliculitis, cutaneous abscesses, cellulitis, wound infection, bacterial myositis) bone and joint (*e.g.*, septic arthritis, osteomyelitis), comprising determining from a sample derived from an individual a increased level of expression of polynucleotide having the sequence of the Sequence Listing. Increased or decreased expression of a polynucleotide of the invention can be measured using any one of the methods well known in the art for the quantitation of polynucleotides, such as, for example, amplification, PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods.

In addition, a diagnostic assay in accordance with the invention for detecting over-expression of a polypeptide of the invention compared to normal control tissue samples may be used to detect the presence of an infection, for example. Assay techniques that can be used to determine levels of a protein, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, West-



ern Blot analysis and ELISA assays.

## Antibodies

The polypeptides of the invention or variants thereof, or cells expressing them can be used as an immunogen to produce antibodies immunospecific for such polypeptides. "Antibodies" as used herein includes monoclonal and polyclonal antibodies, chimeric, single chain, simianized antibodies and humanized antibodies, as well as Fab fragments, including the products of an Fab immunoglobulin expression library.

Antibodies generated against the polypeptides of the invention can be obtained by administering the polypeptides or epitope-bearing fragments, analogues or cells to an animal, preferably a nonhuman, using routine protocols. For preparation of monoclonal antibodies, any technique known in the art that provides antibodies produced by continuous cell line cultures can be used. Examples include various techniques, such as those in Kohler, G. and Milstein, C., *Nature* 256: 495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et al.*, pg. 77-96 in *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc. (1985).

Techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies.

Alternatively phage display technology may be utilized to select antibody genes with binding activities towards the polypeptide either from repertoires of PCR amplified v-genes of lymphocytes from humans screened for possessing recognition of a polypeptide of the invention or from naive libraries (McCafferty, J. *et al.*, (1990), *Nature* 348, 552-554; Marks, J. *et al.*, (1992) *Biotechnology* 10, 779-783). The affinity of these antibodies can also be improved by chain shuffling (Clackson, T. *et al.*, (1991) *Nature* 352, 624-628).

If two antigen binding domains are present each domain may be directed against a different epitope - termed 'bispecific' antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptides to purify the polypeptides by affinity chromatography.

Thus, among others, antibodies against a polypeptide of the invention may be employed to treat infections, particularly bacterial infections and especially infections of the upper respiratory tract (e.g., otitis media, bacterial tracheitis, acute epiglottitis, thyroiditis), lower respiratory (e.g., empyema, lung abscess), cardiac (e.g., infective endocarditis), gastrointestinal (e.g., secretory diarrhoea, splenic abscess, retroperitoneal abscess), CNS (e.g., cerebral abscess), eye (e.g., blepharitis, conjunctivitis, keratitis, endophthalmitis, preseptal and orbital cellulitis, dacryocystitis), kidney and urinary tract (e.g., epididymitis, intrarenal and perinephric abscess, toxic shock syndrome), skin (e.g., impetigo, folliculitis, cutaneous abscesses, cellulitis, wound infection, bacterial myositis) bone and joint (e.g., septic arthritis, osteomyelitis).

Polypeptide variants include antigenically, epitopically or immunologically equivalent variants that form a particular aspect of this invention. The term "antigenically equivalent derivative" as used herein encompasses a polypeptide or its equivalent which will be specifically recognized by certain antibodies which, when raised to the protein or polypeptide according to the invention, interfere with the immediate physical interaction between pathogen and mammalian host. The term "immunologically equivalent derivative" as used herein encompasses a peptide or its equivalent which when used in a suitable formulation to raise antibodies in a vertebrate, the antibodies act to interfere with the immediate physical interaction between pathogen and mammalian host.

The polypeptide, such as an antigenically or immunologically equivalent derivative or a fusion protein thereof is used as an antigen to immunize a mouse or other animal such as a rat or chicken. The fusion protein may provide stability to the polypeptide. The antigen may be associated, for example by conjugation, with an immunogenic carrier protein for example bovine serum albumin (BSA) or keyhole limpet haemocyanin (KLH). Alternatively a multiple antigenic peptide comprising multiple copies of the protein or polypeptide, or an antigenically or immunologically equivalent polypeptide thereof may be sufficiently antigenic to improve immunogenicity so as to obviate the use of a carrier.

Preferably, the antibody or variant thereof is modified to make it less immunogenic in the individual. For example, if the individual is human the antibody may most preferably be "humanized"; where the complementarity determining region(s) of the hybridoma-derived antibody has been transplanted into a human monoclonal antibody, for example as described in Jones, P. *et al.* (1986), *Nature* 321, 522-525 or Tempest *et al.*, (1991) *Biotechnology* 9, 266-273.

The use of a polynucleotide of the invention in genetic immunization will preferably employ a suitable delivery method such as direct injection of plasmid DNA into muscles (Wolff *et al.*, *Hum Mol Genet* 1992, 1:363; Manthorpe *et al.*, *Hum. Gene Ther.* 1993:4, 419), delivery of DNA complexed with specific protein carriers (Wu *et al.*, *J Biol Chem.* 1989: 264,16985), coprecipitation of DNA with calcium phosphate (Benvenisty & Reshef, *PNAS*, 1986:83,9551), encapsulation of DNA in various forms of liposomes (Kaneda *et al.*, *Science* 1989:243,375), particle bombardment (Tang *et al.*, *Nature* 1992, 356:152; Eisenbraun *et al.*, *DNA Cell Biol* 1993, 12:791) and *in vivo* infection using cloned retroviral vectors (Seeger *et al.*, *PNAS* 1984:81,5849).

**Antagonists and agonists - assays and molecules**

Polypeptides of the invention may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics. See, e.g., Coligan *et al.*, *Current Protocols in Immunology* 1(2): Chapter 5 (1991).

The invention also provides a method of screening compounds to identify those which enhance (agonist) or block (antagonist) the action of a polypeptides or polynucleotides of the invention, particularly those compounds that are bacteriostatic and/or bacteriocidal. The method of screening may involve high-throughput techniques. For example, to screen for agonists or antagonists, a synthetic reaction mix, a cellular compartment, such as a membrane, cell envelope or cell wall, or a preparation of any thereof, comprising a polypeptide of the invention and a labeled substrate or ligand of such polypeptide is incubated in the absence or the presence of a candidate molecule that may be an agonist or antagonist of a polypeptide of the invention. The ability of the candidate molecule to agonize or antagonize a polypeptide of the invention is reflected in decreased binding of the labeled ligand or decreased production of product from such substrate. Molecules that bind gratuitously, *i.e.*, without inducing the effects of a polypeptide of the invention are most likely to be good antagonists. Molecules that bind well and increase the rate of product production from substrate are agonists. Detection of the rate or level of production of product from substrate may be enhanced by using a reporter system. Reporter systems that may be useful in this regard include but are not limited to colorimetric labeled substrate converted into product, a reporter gene that is responsive to changes in polynucleotide or polypeptide activity, and binding assays known in the art.

Another example of an assay for antagonists of polypeptides of the invention is a competitive assay that combines any such polypeptide and a potential antagonist with a compound which binds such polypeptide, natural substrates or ligands, or substrate or ligand mimetics, under appropriate conditions for a competitive inhibition assay. A polypeptide of the invention can be labeled, such as by radioactivity or a colorimetric compound, such that the number of such polypeptide molecules bound to a binding molecule or converted to product can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to a polynucleotide or polypeptide of the invention and thereby inhibit or extinguish its activity. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a binding molecule, without inducing activities induced by a polypeptide of the invention, thereby preventing the action of such polypeptide by excluding it from binding.

Potential antagonists include a small molecule that binds to and occupies the binding site of the polypeptide thereby preventing binding to cellular binding molecules, such that normal biological activity is prevented. Examples of small molecules include but are not limited to small organic molecules, peptides or peptide-like molecules. Other potential antagonists include antisense molecules (see Okano, *J. Neurochem.* 56: 560 (1991); *OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION*, CRC Press, Boca Raton, FL (1988), for a description of these molecules). Preferred potential antagonists include compounds related to and variants of a polypeptide of the invention.

Each of the DNA sequences provided herein may be used in the discovery and development of antibacterial compounds. The encoded protein, upon expression, can be used as a target for the screening of antibacterial drugs. Additionally, the DNA sequences encoding the amino terminal regions of the encoded protein or Shine-Delgarno or other translation facilitating sequences of the respective mRNA can be used to construct antisense sequences to control the expression of the coding sequence of interest.

The invention also provides the use of the polypeptide, polynucleotide or inhibitor of the invention to interfere with the initial physical interaction between a pathogen and mammalian host responsible for sequelae of infection. In particular the molecules of the invention may be used: in the prevention of adhesion of bacteria, in particular gram positive bacteria, to mammalian extracellular matrix proteins on in-dwelling devices or to extracellular matrix proteins in wounds; to block protein-mediated mammalian cell invasion by, for example, initiating phosphorylation of mammalian tyrosine kinases (Rosenshine *et al.*, *Infect. Immun.* 60:2211 (1992); to block bacterial adhesion between mammalian extracellular matrix proteins and bacterial proteins that mediate tissue damage and; to block the normal progression of pathogenesis in infections initiated other than by the implantation of in-dwelling devices or by other surgical techniques.

The antagonists and agonists of the invention may be employed, for instance, to inhibit and treat disease, such as, infections of the upper respiratory tract (e.g., otitis media, bacterial tracheitis, acute epiglottitis, thyroiditis), lower respiratory (e.g., empyema, lung abscess), cardiac (e.g., infective endocarditis), gastrointestinal (e.g., secretory diarrhoea, splenic abscess, retroperitoneal abscess), CNS (e.g., cerebral abscess), eye (e.g., blepharitis, conjunctivitis, keratitis, endophthalmitis, preseptal and orbital cellulitis, dacryocystitis), kidney and urinary tract (e.g., epididymitis, intrarenal and perinephric abscess, toxic shock syndrome), skin (e.g., impetigo, folliculitis, cutaneous abscesses, cellulitis, wound infection, bacterial myositis) bone and joint (e.g., septic arthritis, osteomyelitis).

*Helicobacter pylori* (herein *H. pylori*) bacteria infect the stomachs of over one-third of the world's population causing

stomach cancer, ulcers, and gastritis (International Agency for Research on Cancer (1994) Schistosomes, Liver Flukes and Helicobacter Pylori (International Agency for Research on Cancer, Lyon, France; <http://www.uicc.ch/ecp/ecp2904.htm>). Moreover, the international Agency for Research on Cancer recently recognized a cause-and-effect relationship between *H. pylori* and gastric adenocarcinoma, classifying the bacterium as a Group I (definite) carcinogen. Preferred antimicrobial compounds of the invention found using screens provided by the invention, particularly broad-spectrum antibiotics, should be useful in the treatment of *H. pylori* infection. Such treatment should decrease the advent of *H. pylori*-induced cancers, such as gastrointestinal carcinoma. Such treatment should also cure gastric ulcers and gastritis.

## Vaccines

Another aspect of the invention relates to a method for inducing an immunological response in an individual, particularly a mammal which comprises inoculating the individual with a polypeptide of the invention, or a fragment or variant thereof, adequate to produce antibody and/or T cell immune response to protect said individual from infection, particularly bacterial infection and most particularly *Staphylococcus aureus* infection. Also provided are methods whereby such immunological response slows bacterial replication. Yet another aspect of the invention relates to a method of inducing immunological response in an individual which comprises delivering to such individual a nucleic acid vector to direct expression of a polynucleotide or polypeptide of the invention, or a fragment or a variant thereof, for expressing such polynucleotide or polypeptide, or a fragment or a variant thereof *in vivo* in order to induce an immunological response, such as, to produce antibody and/or T cell immune response, including, for example, cytokine-producing T cells or cytotoxic T cells, to protect said individual from disease, whether that disease is already established within the individual or not. One way of administering the gene is by accelerating it into the desired cells as a coating on particles or otherwise.

Such nucleic acid vector may comprise DNA, RNA, a modified nucleic acid, or a DNA/RNA hybrid.

A further aspect of the invention relates to an immunological composition which, when introduced into an individual capable or having induced within it an immunological response, induces an immunological response in such individual to a polynucleotide of the invention or protein coded therefrom, wherein the composition comprises a recombinant polynucleotide or protein coded therefrom comprising DNA which codes for and expresses an antigen of said polynucleotide or protein coded therefrom. The immunological response may be used therapeutically or prophylactically and may take the form of antibody immunity or cellular immunity such as that arising from CTL or CD4+ T cells.

A polypeptide of the invention or a fragment thereof may be fused with co-protein which may not by itself produce antibodies, but is capable of stabilizing the first protein and producing a fused protein which will have immunogenic and protective properties. Thus fused recombinant protein, preferably further comprises an antigenic co-protein, such as lipoprotein D from *Hemophilus influenzae*, Glutathione-S-transferase (GST) or beta-galactosidase, relatively large co-proteins which solubilize the protein and facilitate production and purification thereof. Moreover, the co-protein may act as an adjuvant in the sense of providing a generalized stimulation of the immune system. The co-protein may be attached to either the amino or carboxy terminus of the first protein.

Provided by this invention are compositions, particularly vaccine compositions, and methods comprising the polypeptides or polynucleotides of the invention and immunostimulatory DNA sequences, such as those described in Sato, Y. *et al.* Science 273: 352 (1996).

Also, provided by this invention are methods using the described polynucleotide or particular fragments thereof which have been shown to encode non-variable regions of bacterial cell surface proteins in DNA constructs used in such genetic immunization experiments in animal models of infection with *Staphylococcus aureus* will be particularly useful for identifying protein epitopes able to provoke a prophylactic or therapeutic immune response. It is believed that this approach will allow for the subsequent preparation of monoclonal antibodies of particular value from the requisite organ of the animal successfully resisting or clearing infection for the development of prophylactic agents or therapeutic treatments of bacterial infection, particularly *Staphylococcus aureus* infection, in mammals, particularly, humans.

The polypeptide may be used as an antigen for vaccination of a host to produce specific antibodies which protect against invasion of bacteria, for example by blocking adherence of bacteria to damaged tissue. Examples of tissue damage include wounds in skin or connective tissue caused, e.g., by mechanical, chemical or thermal damage or by implantation of indwelling devices, or wounds in the mucous membranes, such as the mouth, mammary glands, urethra or vagina.

The invention also includes a vaccine formulation which comprises an immunogenic recombinant protein of the invention together with a suitable carrier. Since the protein may be broken down in the stomach, it is preferably administered parenterally, including, for example, administration that is subcutaneous, intramuscular, intravenous, or intradermal. Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the bodily

fluid, preferably the blood, of the individual; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

While the invention has been described with reference to certain protein, such as, for example, those set forth in the Sequence Listing, it is to be understood that this covers fragments of the naturally occurring protein and similar proteins with additions, deletions or substitutions which do not substantially affect the immunogenic properties of the recombinant protein.

#### Compositions, kits and administration

The invention also relates to compositions comprising the polynucleotide or the polypeptides discussed above or their agonists or antagonists. The polypeptides of the invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to a subject. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of a polypeptide of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration. The invention further relates to diagnostic and pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention.

Polypeptides and other compounds of the invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes among others.

In therapy or as a prophylactic, the active agent may be administered to an individual as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic.

Alternatively the composition may be formulated for topical application for example in the form of ointments, creams, lotions, eye ointments, eye drops, ear drops, mouthwash, impregnated dressings and sutures and aerosols, and may contain appropriate conventional additives, including, for example, preservatives, solvents to assist drug penetration, and emollients in ointments and creams. Such topical formulations may also contain compatible conventional carriers, for example cream or ointment bases, and ethanol or oleyl alcohol for lotions. Such carriers may constitute from about 1% to about 98% by weight of the formulation; more usually they will constitute up to about 80% by weight of the formulation.

For administration to mammals, and particularly humans, it is expected that the daily dosage level of the active agent will be from 0.01 mg/kg to 10 mg/kg, typically around 1 mg/kg. The physician in any event will determine the actual dosage which will be most suitable for an individual and will vary with the age, weight and response of the particular individual. The above dosages are exemplary of the average case. There can, of course, be individual instances where higher or lower dosage ranges are merited, and such are within the scope of this invention.

In-dwelling devices include surgical implants, prosthetic devices and catheters, i.e., devices that are introduced to the body of an individual and remain in position for an extended time. Such devices include, for example, artificial joints, heart valves, pacemakers, vascular grafts, vascular catheters, cerebrospinal fluid shunts, urinary catheters, continuous ambulatory peritoneal dialysis (CAPD) catheters.

The composition of the invention may be administered by injection to achieve a systemic effect against relevant bacteria shortly before insertion of an in-dwelling device. Treatment may be continued after surgery during the in-body time of the device. In addition, the composition could also be used to broaden perioperative cover for any surgical technique to prevent bacterial wound infections, especially *Staphylococcus aureus* wound infections.

Many orthopedic surgeons consider that humans with prosthetic joints should be considered for antibiotic prophylaxis before dental treatment that could produce a bacteremia. Late deep infection is a serious complication sometimes leading to loss of the prosthetic joint and is accompanied by significant morbidity and mortality. It may therefore be possible to extend the use of the active agent as a replacement for prophylactic antibiotics in this situation.

In addition to the therapy described above, the compositions of this invention may be used generally as a wound treatment agent to prevent adhesion of bacteria to matrix proteins exposed in wound tissue and for prophylactic use in dental treatment as an alternative to, or in conjunction with, antibiotic prophylaxis.

Alternatively, the composition of the invention may be used to bathe an indwelling device immediately before insertion. The active agent will preferably be present at a concentration of 1 µg/ml to 10mg/ml for bathing of wounds or

indwelling devices.

A vaccine composition is conveniently in injectable form. Conventional adjuvants may be employed to enhance the immune response. A suitable unit dose for vaccination is 0.5-5 microgram/kg of antigen, and such dose is preferably administered 1-3 times and with an interval of 1-3 weeks. With the indicated dose range, no adverse toxicological effects will be observed with the compounds of the invention which would preclude their administration to suitable individuals.

Each reference disclosed herein is incorporated by reference herein in its entirety. Any patent application to which this application claims priority is also incorporated by reference herein in its entirety.

## TABLES

Certain pertinent data for each of the polypeptides and polynucleotides set forth in the Sequence Listing are summarized in the following Table.

### TABLE 1

Provided in this Table is the closest homologue of each polypeptide encoded by each ORF of the invention based on a comparison of the sequences of in the Sequence Listing with sequences available in the public domain (see the right hand column entitled "Desc"). Where no significant homologue was detected "unknown" appears in the column. Preferred polypeptides encoded by the ORFs of the invention, particularly full length proteins either obtained using such ORFs or encoded entirely by such ORFs, are ones that have a biological function of the homologue listed, among other functions. The analysis used to determine each homologue listed in Table 1 was either BlastP, BlastX or MP-Search, each of which is well known. In the left hand column the SEQ ID NO: of each DNA sequence in the Sequence Listing appears. In the center column, the SEQ ID NO: of each protein (polypeptide) sequence of each sequence in the Sequence Listing appears across from the DNA sequence which encodes it (which DNA sequence is in the left hand column). In some instances the DNA sequence encodes more than one protein sequence and so is listed more than once, each listing being shown next to the encoded protein sequence.

**TABLE 1**

	<b>DNA No.:</b>	<b>Protein No.:</b>	<b>Assembly ID</b>	<b>Description</b>
5	1.	260.	2698808	Unknown
	2.	261.	2700024	Unknown
	3.	262.	2700052	Probable ABC Transporter
10	4.	263.	2700068	Unknown
	5.	264.	2700110	Unknown
	6.	265.	2700242	Polyribo-Nucleotide Nucleotidyl-Transferase (Ec 2.7.7.8) (Polynucleotide Phosphor Ylase) (Pnpase). - Escherichia Coli.
15	7.	266.	2700302	Unknown
	8.	267.	2700520	Unknown
	9.	268.	2700548	Unknown
20	10.	269.	2700738	Unknown
	11.	270.	2700754	Unknown
	12.	271.	2701136	Excinuclease ABC Subunit B (Dina Protein) (Fragment). - Bacillus Subtilis.
25	13.	272.	2701728	Unknown
	14.	273.	2702000	Unknown
	15.	274.	2702042	2-Oxoglutarate Dehydrogenase E1 Component (Ec 1.2.4.2) (Alpha- Ketoglutarate De- Hydrogenase). - Escherichia Coli.
30	16.	275.	2702062	Unknown
	17.	276.	2702076	Unknown
35	18.	277.	2702336	Unknown
	19.	278.	2702342	Unknown
	20.	279.	2702482	Ribonuclease P Protein Component (Ec 3.1.26.5) (Protein C5) (Rnase P). - Bacill Us Subtilis.
40	21.	280.	2702500	Riboflavin Synthase Alpha Chain (Ec 2.5.1.9). - Bacillus Subtilis.
45	22.	281.	2702520	ECKBLTDH NCBI Gi: 41862 - Escherichia Coli.
	23.	282.	2702580	Glucose Inhibited Division Protein B. - Bacillus Subtilis.
50	24.	283.	2702648	Unknown

55

EP 0 841 394 A2

	DNA No.:	Protein No.:	Assembly ID	Description
5	25.	284.	2702700	Flagellar Hook-Associated Protein 1 (Hap1). - Salmonella Typhimurium.
	26.	285.	2702838	Pyrimidine Nucleoside Transport Protein. - Bacillus Subtilis.
10	27.	286.	2702934	Protein Dfp. - Escherichia Coli.
	28.	287.	2703752	Unknown
	29.	288.	2703916	Diaminopimelate Decarboxylase (Ec 4.1.1.20) (Dap Decarboxylase). - Bacillus Sub Tilis.
15	30.	289.	2704026	Unknown
	31.	290.	2704028	Unknown
	32.	291.	2704448	Unknown
20	33.	292.	2704502	Unknown
	34.	293.	2704680	Unknown
	34.	294.	2704680	Unknown
	35.	295.	2704892	Unknown
25	36.	296.	2705066	Unknown
	37.	297.	2705204	Unknown
	38.	298.	2705378	Unknown
	39.	299.	2705414	Unknown
30	40.	300.	2705440	Methylated-DNA-Protein-Cysteine Methyltransferase (Dat1) Homolog - Haemophilus Influenzae (Strain Rd KW20)
35	41.	301.	2705526	Ornithine Carbamoyltransferase Chain I (Ec 2.1.3.3) (Otcase-1) (Fragment). - Sa Lmonella Typhimurium.
	42.	302.	2705632	Unknown
40	43.	303.	2705646	Unknown
	44.	304.	2705712	Ribonucleoside-Diphosphate Reductase 2 Alpha Chain (Ec 1.17.4.1) (Ribonucleotid E Reductase 2). - Salmonella Typhimurium.
45	45.	305.	2705740	Unknown
	46.	306.	2705742	Unknown
	47.	307.	2705876	D-Lactate Dehydrogenase (Ec 1.1.1.28) (D- Ldh). - Lactobacillus Plantarum.
50	48.	308.	2706108	Unknown
	49.	309.	2706282	Unknown
	50.	310.	2706344	Unknown

55

	DNA No.:	Protein No.:	Assembly ID	Description
5	51.	311.	2706452	Nickel-Binding Periplasmic Protein Precursor. - Escherichia Coli.
	52.	312.	2706548	Stringent Response-Like Protein - Streptococcus Equisimilis
10	53.	313.	2706730	Unknown
	54.	314.	2706734	Phosphate Regulatory Protein - Rhizobium Meliloti
15	55.	315.	2706832	Diphtheria Toxin Repressor (Iron-Dependent Diphtheria Tox Regulatory Element) (Tox Regulatory Factor). - Corynebacterium Diphtheriae.
20	56.	316.	2706930	Unknown
	57.	317.	2707190	Unknown
	58.	318.	2707592	Phosphoribosylaminoimidazole Carboxylase Catalytic Subunit (Ec 4.1.1.21) (Air C Arboxylase) (Airc). - Bacillus Subtilis.
25	59.	319.	2707978	Unknown
	60.	320.	2708140	Unknown
	61.	321.	2708372	Unknown
30	62.	322.	2708478	Unknown
	63.	323.	2708862	Nitrite Reductase (Nad(P)H) (Ec 1.6.6.4). - Escherichia Coli.
	64.	324.	2709034	Unknown
35	65.	325.	2709098	Unknown
	66.	326.	2709168	Unknown
	67.	327.	2709352	Unknown
40	68.	328.	2709366	Rep827 Protein - Staphylococcus Sp.
	69.	329.	2709474	Unknown
	70.	330.	2709570	Unknown
	70.	331.	2709570	Unknown
45	71.	332.	2709600	Unknown
	72.	333.	2709642	Unknown
	72.	334.	2709642	Unknown
	73.	335.	2709736	Unknown
50	74.	336.	2709760	Unknown
	75.	337.	2709766	Unknown
	76.	338.	2709772	Unknown

55



EP 0 841 394 A2

	DNA No.:	Protein No.:	Assembly ID	Description
5	77.	339.	2709790	Unknown
	78.	340.	2709830	Udp-N-Acetylmuramoylalanine--D-Glutamate Ligase (Ec 6.3.2.9) (Udp-N-Acetylmuramoyl-L-Alanyl-D-Glutamate Synthetase). - Bacillus Subtilis.
10	79.	341.	2709848	Unknown
	80.	342.	2709852	Unknown
	81.	343.	2709888	Unknown
15	82.	344.	2709954	Unknown
	83.	345.	2709992	Unknown
	84.	346.	2709994	Heat Shock Protein Hslu. - Bacillus Subtilis.
20	85.	347.	2710014	Tryptophan Synthase (Ec 4.2.1.20). - Neurospora Crassa.
	86.	348.	2710028	Single-Strand Binding Protein (Ssb) (Helix-Destabilizing Protein). - Bacillus S Ubtilis.
25	87.	349.	2710092	Elongation Factor Tu (Ef-Tu). - Bacillus Subtilis.
	88.	350.	2710096	Unknown
	89.	351.	2710272	Unknown
30	90.	352.	2710276	Unknown
	91.	353.	2710330	Unknown
	92.	354.	2943530	Unknown
	93.	355.	2943556	Unknown
35	94.	356.	2943564	Unknown
	95.	357.	2943624	Possible Sodium-Dependent Phosphate Transporter
	96.	358.	2943704	Unknown
40	96.	359.	2943704	Unknown
	97.	360.	2943710	Anaerobic Ribonucleoside-Triphosphate Reductase (Ec 1.17.4.2). - Escherichia Co Li.
45	98.	361.	2943716	Unknown
	99.	362.	2943742	U00013 NCBI Gi: 466868NCBI Gi: - Mycobacterium Leprae.
	100.	363.	2943746	Unknown
50	101.	364.	2943774	Unknown
	102.	365.	2943782	Unknown
	103.	366.	2943806	Unknown
	104.	367.	2943822	Unknown
55	105.	368.	2943834	Unknown

EP 0 841 394 A2

	DNA No.:	Protein No.:	Assembly ID	Description
5	106.	369.	2943864	Unknown
	107.	370.	2943956	Unknown
	108.	371.	2943960	Unknown
	108.	372.	2943960	Unknown
10	109.	373.	2944036	Deoxyribose-Phosphate Aldolase (EC 4.1.2.4) - Mycoplasma Hominis (SGC3)
	110.	374.	2944044	Unknown
15	110.	375.	2944044	Biotin Synthetase (Ec 2.8.1.-). - Bacillus Sphaericus.
	111.	376.	2944066	Unknown
	112.	377.	2944114	Unknown
20	113.	378.	2944126	Sirohem Synthase (Contains: Uroporphyrin- Iii C-Methyltransferase (Ec 2.1.1.107) (Urogen Iii Methylase) (Sumt) (Uroporphyrinogen Iii Methylase) / Precorrin-2 Oxidase (Ec 1.-.-.-) / 25 Ferrocyclase (Ec 4.99.1.-)). - Escherichia Coli.
30	114.	379.	2944146	Mercuric Reductase (Ec 1.16.1.1) (Hg(Ii) Reductase). - Staphylococcus Aureus.
	115.	380.	2944210	Unknown
	116.	381.	2944212	Unknown
35	117.	382.	2944262	Regulatory Protein Pfor - Clostridium Perfringens
	118.	383.	2944264	Replicative Dna Helicase (Ec 3.6.1.-). - Bacillus Subtilis.
40	119.	384.	2944276	Ribokinase (Rbsk) Homolog - Haemophilus Influenzae (Strain Rd KW20)
	120.	385.	2944306	Dna-Invertase Hin. - Salmonella Typhimurium.
45	121.	386.	2944308	Unknown
	122.	387.	2944310	Histidine Ammonia-Lyase (Ec 4.3.1.3) (Histidase). - Bacillus Subtilis.
	123.	388.	2944342	Unknown
50	124.	389.	2944344	Primosomal Protein N' (Replication Factor Y). - Escherichia Coli.
	125.	390.	2944360	Unknown
55	126.	391.	2944364	Unknown

	DNA No.:	Protein No.:	Assembly ID	Description
5	127.	392.	2944366	Starvation Sensing Protein Rspb (Ec 1.1.1.). - Escherichia Coli.
	128.	393.	2944390	Probable ABC Transporter
10	128.	394.	2944390	Atp-Binding Protein Abc. - Escherichia Coli.
	129.	395.	2944414	Potassium-Transporting Atpase (Ec 3.6.1.36), A Chain (Atp Phosphohydrolase (Potassium-Transporting), A Chain). - Escherichia Coli.
15	130.	396.	2944426	Unknown
	131.	397.	2944434	Unknown
	132.	398.	2944446	Unknown
20	133.	399.	2944452	Unknown
	134.	400.	2944464	Lipoprotein Nlpd Precursor. - Escherichia Coli.
25	135.	401.	2944492	Pts System, Galacticol-Specific Iia Component (Eiia-Gat) (Galacticol- Permease Iia Component) (Phosphotransferase Enzyme Ii, A Component) (Ec 2.7.1.69). - E Schierichia Coli.
30	136.	402.	2944508	Unknown
	137.	403.	2944522	1-Phosphofructokinase (Fruk) Homolog - Haemophilus Influenzae (Strain Rd KW20)
35	138.	404.	2944556	Folypolyglutamate Synthase (Ec 6.3.2.17) (Folypoly-Gamma-Glutamate Synthetase) (Fpgs). - Bacillus Subtilis.
40	139.	405.	2944576	Lactam Utilization Protein (Lamb) Homolog - Haemophilus Influenzae (Strain Rd K W20)
	140.	406.	2944578	Lactam Utilization Protein (Lamb) Homolog - Haemophilus Influenzae (Strain Rd K W20)
45	141.	407.	2944584	Possible Glutathione Peroxidase
	142.	408.	2944592	BK5TATP NCB - Bacteriophage BK5-T DNA.
50	143.	409.	2944644	Phosphoribosylformylglycinamide Cyclo- Ligase (Ec 6.3.3.1) (Airs) (Phosphoribo Syl-Aminoimidazole Synthetase) (Air Synthase). - Bacillus Subtilis.
55	144.	410.	2944648	Unknown

	DNA	Protein	Assembly	Description
	No.:	No.:	ID	
5	145.	411.	2944654	Unknown
	146.	412.	2944658	Indole-3-Glycerol Phosphate Synthase (Ec 4.1.1.48) (Igps). - Lactococcus Lactis (Subsp. Lactis) (Streptococcus Lactis).
10	147.	413.	2944670	Unknown
	148.	414.	2944694	Unknown
	149.	415.	2944706	Unknown
15	150.	416.	2944720	Unknown
	151.	417.	2944746	Unknown
	152.	418.	2944760	Unknown
20	153.	419.	2944782	Pts System, Sucrose-Specific Iiabc Component (Eiiabc-Scr) (Sucrose- Permease Ii Abc Component) (Phosphotransferase Enzyme Ii, Abc Component) (Ec 2.7.1.69) (E Ii-Scr). - Streptococcus Mutans.
25	154.	420.	2944800	Right Origin-Binding Protein. - Escherichia Coli.
	155.	421.	2944806	Unknown
	156.	422.	2944808	Unknown
30	157.	423.	2944838	Phospho-N-Acetylmuramoyl-Pentapeptide-Transferase (Ec 2.7.8.13). - Bacillus Sub Tilis.
35	158.	424.	2944882	Oligopeptide Transport Atp-Binding Protein Oppf. - Lactococcus Lactis (Subsp. L Actis) (Streptococcus Lactis).
	159.	425.	2944888	Unknown
40	159.	426.	2944888	Unknown
	160.	427.	2944902	Unknown
	161.	428.	2944930	Unknown
	162.	429.	2944932	Unknown
45	163.	430.	2944936	Unknown
	164.	431.	2944966	30s Ribosomal Protein S4 (Bs4). - Bacillus Subtilis.
	165.	432.	2944988	Unknown
50	166.	433.	2945022	Unknown
	167.	434.	2945040	Unknown
	168.	435.	2945060	Unknown

55

EP 0 841 394 A2

	DNA No.:	Protein No.:	Assembly ID	Description
5	169.	436.	2945070	Unknown
	170.	437.	2945078	Isocitrate Dehydrogenase (Nadp) (Ec 1.1.1.42) (Oxalosuccinate Decarboxylase) (I Dh) (Nadp+-Specific Icdh) (Idp). - Bacillus Subtilis.
10	171.	438.	2945082	Unknown
	172.	439.	2945186	Unknown
	173.	440.	2945250	Unknown
15	174.	441.	2945254	Unknown
	175.	442.	2945260	Unknown
	176.	443.	2945316	Tagd Protein. - Vibrio Cholerae.
20	177.	444.	2945338	4-Methyl-5-(Beta-Hydroxyethyl)Thiazole Monophosphate Synthesis Protein Thif - E Scherichia Coli
	178.	445.	2945378	Pyrroline-5-Carboxylate Reductase (EC 1.5.1.2) - Arabidopsis Thaliana
25	179.	446.	2945474	Unknown
	180.	447.	2945520	Unknown
	181.	448.	2945550	Unknown
30	182.	449.	2945580	Nadh Dehydrogenase I Chain G (Ec 1.6.5.3). - Escherichia Coli.
	183.	450.	2945772	Unknown
	184.	451.	2945780	Signal Peptidase I S (Ec 3.4.21.89) (Spase I) (Leader Peptidase I). - Bacillus Subtilis.
35	184.	452.	2945780	Signal Peptidase I S (Ec 3.4.21.89) (Spase I) (Leader Peptidase I). - Bacillus Subtilis.
	185.	453.	2945792	Fibronectin-Binding Protein Precursor (Fnbp). - Staphylococcus Aureus.
40	186.	454.	2945794	Unknown
	187.	455.	2945802	Rod Shape-Determining Protein. - Escherichia Coli.
45	188.	456.	3038342	X-Pro Dipeptidase (EC 3.4.13.9) - Lactobacillus Delbrueckii
	189.	457.	3038352	Unknown
	190.	458.	3038362	Unknown
50	191.	459.	3038372	Unknown
	192.	460.	3038374	Phosphotransferase System Enzyme II - Staphylococcus Carnosus
55	193.	461.	3038406	Unknown

	DNA No.:	Protein No.:	Assembly ID	Description
5	194.	462.	3038410	CrtD Protein - Rhodobacter Sphaeroides
	195.	463.	3038424	Tetrahydropteroyltryglutamate
				Methyltransferase (Mete) Homolog -
10				Haemophilus In Fluenzae (Strain Rd
				KW20)
	196.	464.	3038426	Unknown
	197.	465.	3038428	Exopolysaccharide Production Protein Pss. -
15				Rhizobium Leguminosarum (Biovar Pha
				Seoli).
	198.	466.	3038438	Galacticol-1-Phosphate Dehydrogenase (Ec
				1.1.1.-). - Escherichia Coli.
20	199.	467.	3038442	Unknown
	200.	468.	3038450	Excinuclease Abc Subunit C. - Bacillus
				Subtilis.
	201.	469.	3038458	Nitrite Reductase (Nad(P)H) Small Subunit
25				(Ec 1.6.6.4). - Bacillus Subtilis.
	202.	470.	3038474	Unknown
	203.	471.	3038482	Probable Imidazoleglycerol-Phosphate
				Dehydratase (Ec 4.2.1.19). - Anabaena Sp.
30				(Strain Pcc 7120).
	204.	472.	3038492	Unknown
	205.	473.	3038496	SXSCRBA Sucrose Repressor -
35				Staphylococcus Xylosus.
	206.	474.	3038498	Nitrate Reductase Alpha Chain - Bacillus
				Subtilis (Fragment)
	206.	475.	3038498	Nitrate Reductase Alpha Chain - Bacillus
40				Subtilis (Fragment)
	207.	476.	3038504	3-Isopropylmalate Dehydratase (Ec 4.2.1.33)
				(Isopropylmalate Isomerase) (Alpha- Ipm
				Isomerase) (Ipmi). - Lactococcus Lactis
45				(Subsp. Lactis) (Streptococcus La Ctis).
	207.	477.	3038504	3-Isopropylmalate Dehydratase (Ec 4.2.1.33)
				(Isopropylmalate Isomerase) (Alpha- Ipm
				Isomerase) (Ipmi). - Lactococcus Lactis
50				(Subsp. Lactis) (Streptococcus La Ctis).
	208.	478.	3038510	Unknown
	209.	479.	3038536	Unknown
55	210.	480.	3038538	Unknown

EP 0 841 394 A2

	DNA No.:	Protein No.:	Assembly ID	Description
5	211.	481.	3038544	Formate Dehydrogenase Alpha Chain (Ec 1.2.1.2). - Methanobacterium Formicicum.
	212.	482.	3038550	Peptide Transport System Permease Protein Sapb. - Salmonella Typhimurium.
10	212.	483.	3038550	Dipeptide Transport System Permease Protein Dppb. - Bacillus Subtilis.
	213.	484.	3038552	Unknown
	214.	485.	3038578	Unknown
15	215.	486.	3038588	Seca Protein - Staphylococcus Carnosus
	216.	487.	3038590	Di-Tripeptide Transporter. - Lactococcus Lactis (Subsp. Lactis) (Streptococcus Lactis).
20	217.	488.	3038594	Histidinol Dehydrogenase (Ec 1.1.1.23) (Hdh). - Lactococcus Lactis (Subsp. Lact Is) (Streptococcus Lactis).
25	218.	489.	3038596	Histidinol Dehydrogenase (Ec 1.1.1.23) (Hdh). - Escherichia Coli.
	219.	490.	3038624	Thiolase (EC 2.3.1.9) - Clostridium Acetobutylicum
30	220.	491.	3038642	Exou Protein - Rhizobium Meliloti
	221.	492.	3038658	Unknown
	222.	493.	3038686	Unknown
	223.	494.	3038692	Unknown
35	224.	495.	3038726	Unknown
	225.	496.	3038746	Unknown
	226.	497.	3038752	Unknown
	227.	498.	3038764	Unknown
40	228.	499.	3038766	Glucokinase Regulator Homolog - Haemophilus Influenzae (Strain Rd KW20)
	229.	500.	3038772	Unknown
	230.	501.	3038798	Bile Acid-Coenzyme A Ligase (Ec 6.-.-.-). - Eubacterium Sp. (Strain Vpi 12708).
45	231.	502.	3038810	Probable Reductase
	232.	503.	3038816	Unknown
	232.	504.	3038816	Unknown
50	233.	505.	3038848	Unknown
	234.	506.	3038858	3-Dehydroquinate Synthase (Ec 4.6.1.3). - Bacillus Subtilis.
	235.	507.	3038872	Unknown
55	236.	508.	3038896	Unknown

	DNA	Protein	Assembly	Description
5	No.:	No.:	ID	
	237.	509.	3038902	Unknown
	238.	510.	3038954	Unknown
	239.	511.	3038956	Unknown
10	240.	512.	3038964	P115 Protein - Mycoplasma Hyorhinis (SGC3)
	241.	513.	3039072	Pyruvate Oxidase (EC 1.2.3.3) Mutant With Pro 178 Replaced By Ser, Ser 188 Repl
15				Aced By Asn, And Ala 458 Replaced By Val (P178s,S188n,A458v), Chain A - Lacto
				Bacillus Plantarum
	242.	514.	3039118	Unknown
20	243.	515.	3039122	Unknown
	243.	516.	3039122	Unknown
	244.	517.	3039124	Unknown
	245.	518.	3039126	Unknown
25	246.	519.	3039132	Glucose-Fructose Oxidoreductase (EC 1.1.-.-) Precursor - Zymomonas Mobilis
	246.	520.	3039132	Unknown
30	247.	521.	3039156	Putative Ornithine Carbamoyltransferase (Ec 2.1.3.3) (Otc) (Hypothetical Protein In Sagg 3'region) (Fragment). - Streptococcus
				Pyogenes.
35	248.	522.	3039170	Unknown
	249.	523.	3039172	Carbamoyl-Phosphate Synthase, Pyrimidine-Specific, Large Chain (Ec 6.3.5.5) (Ca
				Rbamoyl-Phosphate Synthetase Ammonia Chain). - Bacillus Subtilis.
40	250.	524.	3039176	Acetoin(Diacetyl) Reductase (Ec 1.1.1.5) (Acetoin Dehydrogenase) (Ar). - Klebsi Ella
				Terrigena.
45	251.	525.	3039182	Formate Dehydrogenase Alpha Chain (Ec 1.2.1.2). - Methanobacterium Formicicum.
	252.	526.	3039200	Uvr-402 Protein - Streptococcus
				Pneumoniae Plasmid Psb470
50	253.	527.	3039234	Unknown
	254.	528.	3039244	CDP-Ribitol Pyrophosphorylase -
				Haemophilus Influenzae
	255.	529.	3039246	Unknown
55	256.	530.	3039248	Unknown



DNA No.:	Protein No.:	Assembly ID	Description
257.	531.	3039254	Unknown
257.	532.	3039254	Unknown
258.	533.	3039264	Unknown
259.	534.	3039282	Unknown

**Example 1****Isolation of DNA coding for a Novel Protein from *S. Aureus* WCUH 29**

The polynucleotide having the DNA sequence given herein can be obtained from a library of clones of chromosomal DNA of *S.aureus* WCUH 29 in *E.coli*. Libraries may be prepared by routine methods, for example:

**Methods 1 and 2.**

Total cellular DNA is isolated from *Staphylococcus aureus* strain WCUH29 (NCIMB 40771) according to standard procedures and size-fractionated by either of two methods.

**Method 1.**

Total cellular DNA is mechanically sheared by passage through a needle in order to size-fractionate according to standard procedures. DNA fragments of up to 1 kbp in size are rendered blunt by treatment with exonuclease and DNA polymerase, and *EcoRI* linkers added. Fragments are ligated into the vector Lambda ZapII that has been cut with *EcoRI*, the library packaged by standard procedures and *E.coli* infected with the packaged library. The library is amplified by standard procedures.

**Method 2.**

Total cellular DNA is partially hydrolysed with a combination of four restriction enzymes (*RsaI*, *PstI*, *AluI* and *Bsh1235I*) and size-fractionated according to standard procedures. *EcoRI* linkers are ligated to the DNA and the fragments then ligated into the vector Lambda ZapII that have been cut with *EcoRI*, the library packaged by standard procedures, and *E.coli* infected with the packaged library. The library is amplified by standard procedures.

**Example 2****The determination of expression during infection of a gene from *Staphylococcus aureus* WCUH29**

Necrotic fatty tissue from a four day groin infection of *Staphylococcus aureus* WCUH29 in the mouse is efficiently disrupted and processed in the presence of chaotropic agents and RNAase inhibitor to provide a mixture of animal and bacterial RNA. The optimal conditions for disruption and processing to give stable preparations and high yields of bacterial RNA are followed by the use of hybridisation to a radiolabelled oligonucleotide specific to *Staphylococcus aureus* 16S RNA on Northern blots. The RNase free, DNase free, DNA and protein free preparations of RNA obtained are suitable for Reverse Transcription PCR (RT-PCR) using unique primer pairs designed from the sequence of each gene of *Staphylococcus aureus* WCUH29.

**a) Isolation of tissue infected with *Staphylococcal aureus* WCUH29 from a mouse animal model of infection**

10 ml. volumes of sterile nutrient broth (No.2 Oxoid) are seeded with isolated, individual colonies of *Staphylococcus aureus* WCUH29 from an agar culture plate.

The cultures are incubated aerobically (static culture) at 37 degrees C for 16-20 hours. 4 week old mice (female, 18g-22g, strain MFI) are each infected by subcutaneous injection of 0.5ml. of this broth culture of *Staphylococcus aureus* WCUH29 (diluted in broth to approximately 10<sup>8</sup> cfu/ml.) into the anterior, right lower quadrant (groin area). Mice should be monitored regularly during the first 24 hours after infection, then daily until termination of study. Animals

with signs of systemic infection, i.e. lethargy, ruffled appearance, isolation from group, should be monitored closely and if signs progress to moribundancy, the animal should be culled immediately.

Visible external signs of lesion development will be seen 24-48h after infection. Examination of the abdomen of the animal will show the raised outline of the abscess beneath the skin. The localised lesion should remain in the right lower quadrant, but may occasionally spread to the left lower quadrant, and superiorly to the thorax. On occasions, the abscess may rupture through the overlying skin layers. The affected animal should be culled immediately and the tissues sampled if possible. Failure to cull the animal may result in the necrotic skin tissue overlying the abscess being sloughed off, exposing the abdominal muscle wall.

Approximately 96h after infection, animals are killed using carbon dioxide asphyxiation. To minimise delay between death and tissue processing/storage, mice should be killed individually rather than in groups. The dead animal is placed onto its back and the fur swabbed liberally with 70% alcohol. An initial incision using scissors is made through the skin of the abdominal left lower quadrant, travelling superiorly up to, then across the thorax. The incision is completed by cutting inferiorly to the abdominal lower right quadrant. Care should be taken not to penetrate the abdominal wall. Holding the skin flap with forceps, the skin is gently pulled away from the abdomen. The exposed abscess, which covers the peritoneal wall but generally does not penetrate the muscle sheet completely, is excised, taking care not to puncture the viscera.

The abscess/muscle sheet and other infected tissue, such as the necrotic pads of fatty tissue in the abdominal lower right and left quadrants may require cutting in sections, prior to flash-freezing in liquid nitrogen, thereby allowing easier storage in plastic collecting vials.

#### b) Isolation of *Staphylococcus aureus* WCUH29 RNA from infected tissue samples

4-6 infected tissue samples (each approx 0.5-0.7g) in 2ml screw-cap tubes are removed from -80°C storage into a dry ice ethanol bath. In a microbiological safety cabinet the samples are disrupted individually whilst the remaining samples are kept cold in the dry ice ethanol bath. To disrupt the bacteria within the tissue sample 1ml of TRIzol Reagent (Gibco BRL, Life Technologies) is added followed by enough 0.1mm zirconia/silica beads to almost fill the tube, the lid is replaced taking care not to get any beads into the screw thread so as to ensure a good seal and eliminate aerosol generation. The sample is then homogenised in a Mini-BeadBeater Type BX-4 (Biospec Products). Necrotic fatty tissue is treated for 100 seconds at 5000 rpm in order to achieve bacterial lysis. *In vivo* grown bacteria require longer treatment than *in vitro* grown *S. aureus* WCUH29 which are disrupted by a 30 second bead-beat.

After bead-beating the tubes are chilled on ice before opening in a fume-hood as heat generated during disruption may degrade the TRIzol and release cyanide.

200 microlitres of chloroform is then added and the tubes shaken by hand for 15 seconds to ensure complete mixing. After 2-3 minutes at room temperature the tubes are spun down at 12,000 x g, 4 °C for 15 minutes and RNA extraction is then continued according to the method given by the manufacturers of TRIzol Reagent i.e.: - The aqueous phase, approx 0.6 ml, is transferred to a sterile eppendorf tube and 0.5 ml of isopropanol is added. After 10 minutes at room temperature the samples are spun at 12,000 x g, 4 °C for 10 minutes. The supernatant is removed and discarded then the RNA pellet is washed with 1 ml 75% ethanol. A brief vortex is used to mix the sample before centrifuging at 7,500 x g, 4 °C for 5 minutes. The ethanol is removed and the RNA pellet dried under vacuum for no more than 5 minutes. Samples are then resuspended by repeated pipetting in 100 microlitres of DEPC treated water, followed by 5-10 minutes at 55 °C. Finally, after at least 1 minute on ice, 200 units of Rnasin (Promega) is added. RNA preparations are stored at -80 °C for up to one month. For longer term storage the RNA precipitate can be stored at the wash stage of the protocol in 75% ethanol for at least one year at -20°C.

Quality of the RNA isolated is assessed by running samples on 1% agarose gels. 1 x TBE gels stained with ethidium bromide are used to visualise total RNA yields. To demonstrate the isolation of bacterial RNA from the infected tissue 1 x MOPS, 2.2M formaldehyde gels are run and vacuum blotted to Hybond-N (Amersham). The blot is then hybridised with a <sup>32</sup>P labelled oligonucleotide probe specific to 16s rRNA of *S. aureus* (K. Greisen, *et al.*, *J. Clin. Microbiol.* 32 335-351(1994)). An oligonucleotide selected from the group consisting of the polynucleotides of the Sequence Listing is used as a probe. The size of the hybridising band is compared to that of control RNA isolated from *in vitro* grown *S. aureus* WCUH29. Correct sized bacterial 16s rRNA bands can be detected in total RNA samples which show extensive degradation of the mammalian RNA when visualised on TBE gels.

#### c) The removal of DNA from *Staphylococcus aureus* WCUH29 derived RNA

DNA was removed from 73 microlitre samples of RNA by a 15 minute treatment on ice with 3 units of DNaseI, amplification grade (Gibco BRL, Life Technologies) in the buffer supplied with the addition of 200 units of Rnasin (Promega) in a final volume of 90 microlitres.

The DNase was inactivated and removed by treatment with TRIzol LS Reagent (Gibco BRL, Life Technologies)

according to the manufacturers protocol.

DNase treated RNA was resuspended in 73 microlitres of DEPC treated water with the addition of Rnasin as described in Method 1.

**d) The preparation of cDNA from RNA samples derived from infected tissue**

10 microlitre samples of DNase treated RNA are reverse transcribed using a SuperScript Preamplification System for First Strand cDNA Synthesis kit (Gibco BRL, Life Technologies) according to the manufacturers instructions. 1 nanogram of random hexamers is used to prime each reaction. Controls without the addition of SuperScriptII reverse transcriptase are also run. Both +/-RT samples are treated with RNaseH before proceeding to the PCR reaction

**e) The use of PCR to determine the presence of a bacterial cDNA species**

PCR reactions are set up on ice in 0.2ml tubes by adding the following components:

45 microlitres PCR SUPERMIX (Gibco BRL, Life Technologies).

1 microlitre 50mM MgCl<sub>2</sub>, to adjust final concentration to 2.5mM.

1 microlitre PCR primers (optimally 18-25 basepairs designed to possess similar annealing temperatures), each primer at 10mM initial concentration.

2 microlitres cDNA.

PCR reactions are run on a Perkin Elmer GeneAmp PCR System 9600 as follows:

5 minutes at 95°C, then 50 cycles of 30 seconds each at 94°C, 42°C and 72 °C followed by 3 minutes at 72°C and then a hold temperature of 4°C.

10 microlitre aliquots are then run out on 1% 1 x TBE gels stained with ethidium bromide with PCR product sizes estimated by comparison to a 100 bp DNA Ladder (Gibco BRL, Life Technologies).

RT/PCR controls may include +/- reverse transcriptase reactions, 16s rRNA primers or DNA specific primer pairs designed to produce PCR products from non-transcribed *S. aureus* WCUH29 genomic sequences.

To test the efficiency of the primer pairs they are used in DNA PCR with WCUH29 total DNA. PCR reactions are set up and run as described above using approx. 1 microgram of DNA in place of the cDNA and 35 cycles of PCR rather than 50.

Primer pairs which fail to give the predicted sized product in either DNA PCR or RT/PCR (approx 20%) are PCR failures and as such are uninformative. Of those which give the correct size product with DNA PCR three classes are distinguished in RT/PCR:

1. Genes which are not expressed *in vivo* reproducibly fail to give a product in RT/PCR.

2. Genes which are expressed *in vivo* reproducibly give the correct size product in RT/PCR and show a stronger signal in the +RT samples than in the -RT controls.

3. Genes which may be expressed *in vivo* give similar amounts of product in both +/-RT samples.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Black, Michael

Burnham, Martin

Hodgson, John

Knowles, David

Lonetto, Michael

Nicholas, Richard

Pratt, Julie

Reichard, Richard

Rosenberg, Martin

Ward, Judith

(ii) TITLE OF THE INVENTION: Novel Prokaryotic Polynucleotides,  
Polypeptides and Their Uses

(iii) NUMBER OF SEQUENCES: 534

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: SmithKline Beecham Corporation

(B) STREET: 709 Swedeland Road

(C) CITY: King of Prussia

(D) STATE: PA

(E) COUNTRY: USA

(F) ZIP: 19406-0939

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette

(B) COMPUTER: IBM Compatible

(C) OPERATING SYSTEM: DOS

(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:

(B) FILING DATE:

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 60/027,032

(B) FILING DATE: 24-SEP-1996

EP 0 841 394 A2

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Gimmi, Edward R
- (B) REGISTRATION NUMBER: 38,891
- (C) REFERENCE/DOCKET NUMBER: P50549

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 610-270-4478
- (B) TELEFAX: 610-270-5090
- (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 396 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GAATTCCTGA TAATTATTTN GCGATATAAG TCGTCAACGT ACCAATATTA TCAATAGTCA	60
CTTTANATNC ATCACCTGGT TGGAAAAATT TAGGTGGTTG AATACCTGNA CCANCGTCTG	120
AGTGGTGTAC CAGTTGCAAT AATATCTCCC GGATGNAGTG CAACACATCA TGACATCACT	180
ACTATTACAA CAACATATAT CAGAATCAAG TCGCCAGNGT NACCATCTTG TCTAATTTCA	240
TNGTGACCTC NTGTAACAAT ATTACATTTT CCAGGTAATG GCAGTNCGTC TTGAGNAANG	300
ATTTAAGGAC CCATNGGNTN ANCCNNCAGT TAACTNNGT GGTA AAAATG CTTGATCTTG	360
NTCACTGCGG GCTNTGAGAT CAAGTGATAT CGTAAA	396

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 696 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CCCCTCTATT GCTTNTTGA TACGTTTCTG TTCTTTAAAT CCAAGATATA CCGCAAACAA	60
CATAACAAGA ATGAAACCTA TTATTTGAAT AGGTATTAAT CCATACCATA ATTCAATTGAC	120

EP 0 841 394 A2

ACTTTTGGCT	TTTAACACTG	CAGCTACACG	AGCCATTGGA	CCTCCCCAAG	GTNCCATGTT	180
CATAATCGCC	GCGCTTAATG	CTAATAGTAA	AATCAATAAA	TATTTATTCA	TATTTAACGC	240
TTTATATAAA	GGTAATAATG	CAGGAATAGA	AAGCAAAAAT	GTTACCGCAC	CGGCCCCATC	300
TAATTGGGCT	ATTGTGCCAA	TTAAAGCTGT	CATTGCACAG	ACAATGACGA	CATTGCCTCG	360
TGTCATTAAT	ATTAAGCGTT	TGACAAGCGG	CTTGAATAAA	CCACTATCGT	TCATGATGCC	420
AAAGAAAATA	ATGGCANAGA	TAAACATAAT	AACAACGTNG	ATGACNTGAT	CTAACCCCTT	480
AGCAAAAAT	CCAACCAAAT	CTGNACACT	ATATCCTAAA	ATCATTGCCC	CTAAGCAAGG	540
TATGATTGTC	ATACCTACAN	CTGGATTTAT	CTTNNTGGGC	AATGAGTAAN	CCTACAACCTG	600
AAATAATAAT	NATGNGCCCC	ATCACTGTCA	ACCACATATT	ATCACTATTTC	ATTTATTTTC	660
TCCCTCCACT	TNCAATACAA	TTTACATTCA	CCCCC			696

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 588 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCCTTGNC	AA	TTCATNGGGT	TTTCTGTTTA	ATATTTTTCG	GGTTCTAANC	CCACCATATC	60
AAGTGATTCA	TCTACTCTTT	TATCAATATC	TTTTTCTTCC	CACTTTTTC	TTTGTGCCAC		120
TTTGTGCAAA	AATTTCCTTT	GGGCCTCAAA	TGGGGNATAA	TTCCAATCTN	CCTCCAANAC		180
GTANCCAATA	TCCCAACGCA	TTTCGTAATA	CTGGATAATC	ACTTATTGGT	TTATCTTTAA		240
AATAAATATA	ACCTTCACTT	AAGTGAATGA	GTCGATTAAT	CATTTTAAAT	GTCGTAGTTT		300
TTCCACAACC	TGAAGGTCCA	ATTAGCACAA	AAAATTCACC	CTCATTAATA	TTGAAACTAA		360
TGTTATCGAC	AGCAACATGT	TTGCCATAAC	GCTTAGTTAC	ATTTTAAAC	TTAATCACTT		420
TGCCACCTCT	TTTTTCTCA	TAGCATAAAA	CCGAGATTAT	ATGTATGTAT	TCCCTATTTA		480
ACCACGTTTA	TTACAATTTT	CAAATTTAAA	TGATTTATCC	TTGAACTTTT	TTAACAAAAT		540
AATGAATAAT	AGGTAATCTC	CAGTTAAGAA	ATAGTGTAT	TTTACCTT			588

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 500 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

# EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGGGCAGNCT TTANAATTAT TTNAATGATG ATTAGCAATG CATATACGAT TAATTCGATT	60
ATGGNTTATA TCTNNATTGA TTAANTCACT TATATTTTAC AAGTGTGTTG CTTTGGAGCG	120
TCAACGATGC TATCACATAT TACGGCATAT GAAATTTTAT CTGAAATTCG NAAAAAATTA	180
GCTCAAAAAT TAATGCGCCT CCCTTTAGGN GTAGTGGAAT CTAAGAAAAT AGGTGAATTA	240
AAAAATATAT TTGNCGATAA GGTTGAAACA ATAGAATTAC CTTTAGCACA TATGATTCCT	300
GAAGTTATNG GAAACTTACT TGTGGCAGCT GCTATCTTCT TATACATAAT GNTCATNGAT	360
TGGCGTATTG GATAGGTGNC TTANTAGATA ACGATACCAA TNTCAATTTN CGCTTCTAAA	420
AAAGTNATGT CTTGGATTTT AATGAGANAA TTTGCTGGAC AAANTGAAAT CCAATTNTN	480
TTTTTNATTT GNCGNAATGG	500

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 808 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ACCCTGTTGT CCCCNNGGGT CAATAGACCC CGACATACTT TNNGGCTTCT GAAAATAAAA	60
TTTCAGGNNC GCACTTTAAA GATGGACGTC GATGTCGACT AATGTGATGA CNGCCTTTCA	120
TCTTTTAAA AANTCATTAA TTTCCCTTCT TGTTTAANAC GTACATANAA GAAATACGCA	180
TACGGTACTA ATAAAATAGT TGTATATGTT GCGTGTGTTA ATAATAATAC ACCGATTAAT	240
TCAGGAATGA TGTTTAAGAA GTAATTTGGG TGTTTTGTA ATTTTATATA ATCCAGATTT	300
AATAATAGGA TGGTTAGGTA AAATGAATAA TTTTAATGTC CAAATACCAC CTAAAGTTT	360
AATAACCATA AATAACATGA TATAAGCAAA GATTAATATA ACTAAGCCAA TACCATTTGC	420
AAAGCTAAAT GTATCTTTAT TAATAAATGC CTCTACACCA GCCAATACAT AAATTAAAC	480
GTGTGTTATT GCTAAAACT TCGAATTTT AACGCCATAT TCAACTGCAC CGTCTGCTTT	540
TAATTGTTTT GAGTGANTAA TACATATCTT TAAGCTGACA AGTCTGATAC AGAGAAAGAT	600
AAGTAATATA GATAGANTCA TGATGTCCTC CGTCATTATG TCATATGTAT AAGCGTTGAT	660
TTTGACAACA TAAAGTATTT TATAGATAAA ACTTGTGACA TACTATTAAC TATTTATTAA	720
TTTGTAGTACA TAAATATGTT TCTAAGTATG TGTTTATGTT CAGTATTTTG GATAATTTAA	780
TAATTTTAAG GATATTAAGC GCTTACAG	808

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 806 base pairs
- (B) TYPE: nucleic acid

EP 0 841 394 A2

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CGGAATTC	TTTATTCAT	TATATAA	TATATCG	TATGAT	TGTTT	ATTAGTCCAC	60
CATTTCG	ATTGTG	TATAGT	TGTTAAG	TGTAAC	GAAGGG	TATTAAGTTG	120
AATAACT	ACAATA	TGGCCA	ACATAT	CAAG	GTGCGT	GTAATTC	180
CTGGT	ATTCA	NCCAT	ACGGT				240
GCGTTG	TTAG	AGACCC	TAA	TCATA	CTGAA	GGACTGATC	300
ATATAT	CAGA	AATTAT	GGAT				360
GACTAC	GTT	ATAATT	TGAA	GAAATTT	CTA	TCAGAAGGAC	420
AAATTG	TAA	AGCTAAA	AATT				480
NTGTCT	ATAG	ATGATG	AAGG	AAAGCT	TAA	CTATCATTAA	540
NGGATA	AATGA	TTACTT	CAAA				600
AATTAT	GAGC	GTAAGA	AGGA	AAAACA	ATCA	GTATTAGAT	660
AAATC	AGAG	NACAG	AAAAA				720
TATGGG	TNNC	AAACACT	TAA	AAGAAC	GCTT	ACCAATCTTG	780
GATAAA	ACAG	GCAAAG	CGAG				806
CAACTC	GAAA	CNGACT	AAA	GACAA	AGATA	AATCCGNACC	
CGAAA	ATCAA	ACAAAG	GGTC				
TTGAA	ATGAA	AGTTT	CTTAG	ACTATA	AAAAG	AGATTAGGTA	
TCTAT	TAAAT	TTTAT	TAGAT				
ACTAAT	CTCT	TTTTG	TCTAC	GATAAC	GTA	TATGATTGAT	
TCTAT	TTTACA	CGTAC	AAATG				
GTTTA	AGGTG	ACATAT	CCAT	TATCTT	TGTT	AGATAGAATC	
GTTG	ATTGC	AATAT	TGTAT				
GTGG	ATTGT	TTTTTT	TATT	TTTTT	AGAA	ATGAGAACTA	
CAACT	TAAAG	TATTAA	ACGA				
ATTG	CAACTA	TATAAC	CAGA	TAATTG			

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1033 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTACGG	CATA	CCATG	CGAAA	TAATT	GTTGA	TTCTA	ATGCT	ACAAT	CGGTT	GATTG	TTCTC	60
CCGTG	CTTGC	TGAAC	TTCTC	GAGA	AATCTC	AATAT	ACTTG	TGTAA	ATTTG	CCATTT	TTAT	120
AATCC	TCCAT	ATCGT	GATAA	AGTT	GCTGTT	GATCT	AGGTT	GTGC	CTAACT	GTATAT	TTTCG	180
TTTCT	ATCGT	TTTCT	NTGCG	TTAAC	CATAC	CAGCA	ATTAA	TATAT	CTACA	GTAGAC	ATCC	240
CATTT	ANCCA	GCTAT	ATACT	ACTGC	AGCAC	AGAAT	GAATC	GCCT	GACCT	GTAAC	ATCTT	300
TCAC	ACTATG	TGGAT	GGCAT	AACTG	ACTTG	ATGAT	TCTTC	CTCA	CCACTN	CGGATA	ATTA	360
ATTCT	TNCAC	GCCAT	NTGTC	ACAAT	AACAT	TTNTA	CACCT	AAAT	CATTCC	AGCGG	TTTTA	420
GCAG	CTATTT	NTANAT	CAAC	AAGTA	GATTC	TTATTT	TNTAA	NTTTA	AGTAT	GGTTC	TGGCN	480
CANCT	TNANT	CGGG	GATAAT	CCAAT	CAATA	GCATG	TAATG	AATCA	GGCAT	ATTTT	TCATT	540
TTTGG	GGAAG	AAACC	GTGGT	GATAA	CTAAT	TTGAT	TTGAT	GTTTC	GTGGT	ATAGG	CACAT	600



EP 0 841 394 A2

AAGAAGTTTA ATGCCTCTTT GCCTAAATTC AAATCGACAA TAATGCACTT AGCCTTTTTC 660  
 AATAAGTGTG AACGCTTAAT TAAAAATTCA GCGTAATGT AGTCANACAC TTCCATATCT 720  
 5 GCTAAGCCAT ATGTCATGTC GCCTTCTTTA CTAATTAAAG CTGTATATGA ACCTGTACTC 780  
 GCATTTTCAA ATTGTTGAAC ATGANCCCAA ATTCATAAAT GGTGTGGNCA ATCGTTGAAT 840  
 CATTTCCCAN TCACTATCTT GTCCACTAGC AGATAANAAA GCGACCGTTT CGCCAAGCCT 900  
 ACCTAAGTTC TCAGCAATAT TTC'TTGCTAC GCCACCAATA GAGCGTGTG ACGTTACAGG 960  
 10 ATTTGATGTT TCTGCAACTA AATCTTTATG CACATAAAAC TTACGATCTA CATTGCTGCG 1020  
 GCCAATACAA ACG 1033

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1172 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CTTACACAAC AAGTTGTCAN AAATTATTTT TATGATAATG ATCGNNCATT TACTAGAAAA 60  
 30 GTAAAAGAAT TATTTGTAGC TCATCGAGTT GAAAAACAAT ATAATAAGAA CGAAATTTTA 120  
 CGCTTTTATT TAAATAATAT TTACTTTGGG GATAATCAAT ATACGCTTGA GGGCGCAGCA 180  
 AACCATTATG TGGNACACCC GTGAATAAAA ATAGNACAAC AATGTCTCAC ATCACAGNTT 240  
 NACAAAGCGC TATTTTAGCT AGTAAAGTCA ATGCACCTAG CGTATATAAT NTCAATAATA 300  
 35 TGTCAGAGAA TTTCACGCAA CGTGTTAAGC ACGANCTTAG AAAAAATGAA GCAACAACAT 360  
 NATATCAATG ACACACAATA TCAACAGGCT ATGTCACAAC TAAATCGCTA AATACTNGCA 420  
 TTCATGTGTA AGATGCTTGG AACTATAATG TAGAAGGTTT AATTGAAATG TCATAATGAT 480  
 AAAGTTGAAT TGTTTCTGCT TAATAAATTT TACAGATTTG CTGTGAATCA CACTNTCATT 540  
 40 TTAAAAGATG ATAGTGTGAT TTTTGTAG CATCGAAAAA TATTTAAAGT TAAACATTTG 600  
 CTATAGGAAG TGGCTTATTG AATAATAAAA GAGTATTGTA TAAAACTTGA GAAAGTAACG 660  
 GGGTGATCCA GTGCCGAAGA TTAATAAAAT AGAAGTTCAA AAGAAAAACA AAGAACGTCT 720  
 TAATCTTTTT TTAGACGAAC AATTGANAN GGNANTGGG TATAGATATC GATACATTAG 780  
 45 TCAAATTTAA TTTAAAAAAA GGGCAACAAC TTGAAGCTGC TGACATGGNA GAGGTTCAAN 840  
 AGTATGATCA TTATCGCATA GGTTTAAACA AAGCAATCCA NTATTTATCA TATAAAAAGA 900  
 GAACTGAAAA AGAAGTTATT CAATATTTAC AANAAGAAGA GATATCAGAG CAAGCGATTT 960  
 50 CTGANGTGAT TGAATATTGT TATCGCGAAA AGTTANTCGA CCATCAAGAT TATGCGGAAA 1020  
 GTTTACNAAA TACAANGATT CGCACCAGAG ATAATGGACC TANANTTNAT CANCAAAAAC 1080  
 TTTTCAACC TCGGGTATNG NACCACCTAT CATTGANATG TTTACAGANC TTTCTTAGNG 1140  
 AACAAACAGNN CTTGGTTGAT ATTCATCCAA AN 1172

(2) INFORMATION FOR SEQ ID NO:9:

EP 0 841 394 A2

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 399 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GAAGGCTTGT	CGTCCCATTT	AAGCATCACT	CCGTTTGT	TCGCTTAAAA	TCCCTTGCAT	60
CGATGCTAAC	AATTGATCAA	CATCTTTAAA	TTCTTTATAG	ACTGATGCAA	ATCTAACATA	120
TGAAACTTGA	TCAACATGCA	TTAACAAGTT	CATAACGTGT	TCACCTATAT	CTCGTGAAGA	180
CACTTCCGTA	TGACCNTCAT	CTCGTNATTG	GCATNCAACC	TTGNTTGNTA	TGTCTTCAAG	240
TTGTGGATA	TCTAACTGGT	CGGTTCTCAC	AAGAACGCAC	AAGTCCAATA	AGTATTTTNT	300
CTCTGAAAA	CTGNTCTCTT	GCGCCATCTT	TTNNCACAAC	TATAAGCTGA	CTTAACCTCG	360
ATATGTTCAA	ATTGTAGGGG	GAAACCGTGG	TTCCACAA			399

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 632 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ATTTATATAT	AGTAACAAAA	CTAAATTATA	AATGACAATA	TCATATTTTCG	CAATCACTGT	60
TAGTAATTTA	AATATTTTTC	AAAAACGACA	TAATCATCAG	TTGTTCGAT	TTCAACGAAA	120
TTACTATTTA	AACATAATGG	TTTAAATCA	TTATTCCTC	CAAACAGAAT	CATCGTAATC	180
ACATTAAGGT	TGTCATAGTT	TTTCTTAAAG	TAAGTTGGCA	ATTGTCGCAG	CGCACTTGAA	240
CCAATACCAT	TAAAGCGATA	TGATTATCC	ACTTCAATAG	CAGCAATGTT	CATGCGTTCA	300
CCAATCTGCT	CCAATGCGAT	AAAGCCACCT	NTTGGCANCA	GCATCATAAA	NTTAATGATA	360
TNTACATNGG	CCACTTCTTG	AACGACTTTC	TTCTTCCAAG	CTTTTAATGA	CGAACATTGT	420
TTNGATCTTA	AGTCCAAGTA	TTTGGCCGGT	CCTTTGCATT	GGTCCTCTTC	TTAGTCGNTT	480
GATTGAATGA	TCCAGACGTC	CGTTGGTCCA	GATTGGNCAA	TTCCGANTC	TACTACTGNN	540
TTTGNTGATA	AATCAGGAAT	CACCCTGACA	TTTTGGGAAT	GAATCAATNT	TTCCNGGCCA	600
ATTCCAAACC	TGGTTTTTTC	CAATAATCGN	TT			632

(2) INFORMATION FOR SEQ ID NO:11:

EP 0 841 394 A2

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 438 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GATATACAGC	ACTCAGAATT	TTAATGACAA	TTAATGCAGC	AGTTAACACG	ACAACGCCAT	60
TAAATGCTTC	TTTATGCTTC	ATCACTGATT	CTCATGCTTT	CTTCAATGCA	CTTAACTAAA	120
AACTTCAAAC	TATCAAGCCA	TTGATTTTGT	TTCGTTAAAG	TAATTGTCAI	TGCATTATTT	180
TGAACACCAA	CCTTCATTGT	TCTACCTAAA	GGTTGTGTTG	CTTTGAACAG	CACTTCGCCA	240
TCAATATTTT	CAGTGGCTTT	TACAGATAAA	TGAATATCAA	TTATTTTCCC	TTTATCTTTA	300
ATCAACGTGA	TACCTGAATG	TAATGCGTGT	ACTTTTATTT	CCACTATATC	AAGCAAACGT	360
GCTACTTCAA	CAGGNTAATC	CAATTAAACG	ATCAATTANT	TCATCTTTAN	TGTCGATAAT	420
TGGGNCAAAT	TGTTCCAG					438

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 522 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TACCAAAGTT	CTTCTTTAGA	GGCTGGGGAA	TATNTCCNAA	CTACATCACG	ACGCACTCGA	60
AACGGTTCCT	CGTTGGAAAA	TCGATGTCAT	TTGCGTGATA	TTGCACATCT	ACAAGTTTTC	120
TAAGTAATTC	ACTTCTATCC	ATTTCCATAC	CAACTCGAAC	ACTTACTACT	AAATCTTTAT	180
ATTCTTCAGG	ATTACCTAAA	CCATATATAC	AACTTACACT	AGCAATAATA	ATTACATCAT	240
CGCGTTCAAA	TAATGCACTT	GTAGCAGAAT	GTCGNAGNTG	ATCAATTTCA	TCATTGATTG	300
AGGCATCTTT	TTCAATAAAA	GTGTCAGTAG	ACGGTACGTA	TGCCTCAGGT	TGATAATAAT	360
CATAGTAACT	TACAAAGTAT	TCCACCCTGT	TTTCAGGAAA	AAATCTTTTA	AACTCACTAT	420
ATAATTGCCC	TGCTNATGTT	GTGGTAGGTG	CGATAATTAA	CGTTGGTTTC	CCAACTTCTT	480
TAATAACATT	ACTCATCCGN	CAATGTTTTT	CCTCGNGGCA	GG		522

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

# EP 0 841 394 A2

- (A) LENGTH: 520 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```

ATCGCCACCT GTACCATGCA ACAATATTAG CGTTGGTGCA CCTTTNTTGT CCTTCACGAA    60
AAATATGTTC CATCACAAAT CCTCCTCATT CCAATTAATC ATGTTAACGC TTAGTATTAA    120
ATGGTCTTAT TTCTGATTCT ATATACTCTC TTTTATTTTC TAAAAATGGT GGCAATGATA    180
AACCTNCACC TAATGTTTCA TACGGNTCAT CTTCCATAAA TCCTGGTCCA TCTGNTGAAA    240
TTTCTATTAA AATATGCCCC ACACGTGNAT ATAATGCTTC AAAATAGAAA CGATTAACGA    300
TGCCTGAGTT GGTTAATACC TACCTCTTTA TATTNCGGGT GCCCACGCTT CTTATTTGCA    360
TCAGGATCTT TCCACAACGA AATGACACAT GATGTACCTC ACCATAACTT GAGGGNGGTT    420
GNTGGCCTNN ATCATCTTTT ATTAAGATTA CTGGNCACAT TGGCGNCTTC GCAACTTCAA    480
GGTAATGCGA CATATCTTCA AGCGCAATAG GTGGCATAAC    520

```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 397 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```

GAATTCCATC AGTGGCAGAA CTGAAGTATT TANACCAATA ATGTCATCGA CACTAGTTAC    60
TATTATCGTC TTCTTACCAC TTGTGTTTGT ATCAGGTTCA GTAGGCGAAA TGTTTAGACC    120
TTTTGCATTG GCTATTGCAT TTAGTTTATT AGCATCGTTA TTAGTGTCAT TTACTACTCGT    180
NCCAGCTTTG GCAGCAACAC TATTTAAAAA AGGCGTTAAA CGTCGTAATA AACAAACATCA    240
AGAAGGATTA GGTGTTGTGA GTACAACTTA TAAAAANGTA TNGCATTGGG TCACTAAATC    300
ATAAGTGGA TNGGTAATTA TATTAAGTAC ATTAATTTGG GGNTGNAACT ATTGTTTTTG    360
GGGGGACCNA GACTAGNCAC TTAGCTTTTT TCCAGAA    397

```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 444 base pairs
- (B) TYPE: nucleic acid

EP 0 841 394 A2

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

10

AACCATCTTA	CCGATATCAG	TACAGTACGG	CGTAGGAACG	GGCATCCAGC	GGATTAGAGG	60
TGGTGAAACC	AACCTGGTTG	TTGATAACGA	TACGTACCGT	ACCGCCAAC	TCATAACCAC	120
GCGCTTTCGA	CATGTTCAGG	GTTTCCTGAA	CCACGCCCTG	CCCGGTCACT	NTGGGGTCAC	180
CGNGGATGGT	GATTGGNAGN	ACTTTGTTGN	TGCTCGGNTC	ATCAAGTCTG	TCCAGACGGG	240
CACGAACAGA	ACCGATANCT	ACCNGGCTTA	CAATCTCAAN	GTGAGACGGG	TTAAACGCCA	300
GNGCCAGGTG	NACCAGGGCG	CCATCGGTCT	NTAAGTCAGA	CGAGAAGCCC	ATGTGGTATT	360
TCACGTCACC	CNTGCCGAGG	TGTTCTNTAT	GTTTACCCGG	GGGAACCCG	TCGNAACAAN	420
TCTTTGGGGG	TTTTTTACCC	CAGC				444

20

(2) INFORMATION FOR SEQ ID NO:16:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 387 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

35

GAATTCCTG	GAGGGCAATT	TTGGGCTTGC	AATAACAAT	CATCAAAATA	TTGTTGGTAC	60
TATAGGTTTG	ATTAGACTAG	ATAATAATAT	GTCTGCACTT	AAAAAAATGT	TTGTTGATAA	120
AGGTTATAGG	AATTTAAAAA	TAGGTAAAAA	ACTTTTAGAT	AAAGTTATAA	TGACATGTAA	180
AGAACAAAAT	ATCGATGGNA	TNTATTTAGG	AACAATNGAT	AAATNTATNA	GTGCACAGTA	240
TTNTTATNCT	AATAANGGCT	TCCGCGNAAA	TTAAGCGAGG	AGATTTACCA	AGNAGCTTTC	300
CTAAATCAGA	TGTAGATAAT	AGGTTCCATT	ACAGCAAANT	CNAAAGATTA	ANTCGANGCA	360
CTGAATAANT	GNNATGACAN	CCCTCAG				387

45

(2) INFORMATION FOR SEQ ID NO:17:

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 417 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

55

# EP 0 841 394 A2

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```

5      GAATTCCTAG AAGGAGAATT TGTTAGTGGT GTTGCAGAAA AGCAATGGTT GGAAGAAAAA    60
      GGTATGAAG TAGACTTTGA ATTAATGGCT GAAATAAATC CAACAAGATT ATCATTAAAGT    120
10     GATTGGCTTA AAGTCCAAAA CTATAATAAG TAATGAATGG GGTGCAATCA TGATTCAATC    180
      AATGTGGTTT AATTTGCATG TGCAAGATTT AGAAAAGAGC GCACAGTTT ATAAAGCGTT    240
      AGGATTTAAA ATAAATAGAA ACCCACAAAT GTTAGATAAA ATGGTCGGTA TTCAAATAGG    300
      TCAAACAACC GCAATTTTAA TAGAAAACAA GCATTTTCAA AATGTAAGTC AGCAAAGCCT    360
15     TAATACTGAA CCAATGAAG TGATGATTTC TCTAGGTGTG ACACAATGAG AGTTGAC      417

```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 341 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```

30     GAATTCCTAT ATACTACTTT GAAAAGGTGT GAGCTTAATG ACAACTTTTA GTGAAAAAGA    60
      AAAAATTCAA TTACTAGCAG ATATTGTTGA ACTACAAACT GAAAATAATA ATGAAATAGA    120
35     CGTTTGTAAT TATTTAAAAG ATTTATTCGA CAAGTACGAT ATTAAATCTG AAATTTTGAA    180
      AGTTAATGAA CACCGCGCCA ATATCGTTGC AGAAATCGGT AACGGNTCAC CTATACTCGC    240
      ATTGAGTGGT CATATGGATG TTGTTGGATG CAGGAAATCA AGATAATTGG ACTTATCCCC    300
      CTTTNCAACT CACAGAAAAA GCTGGCAAAT TATACGGTCC G                        341

```

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 351 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```

55     GAATTCCTCA GGTATAACAC TTCNCATATC TTTACGAGNA ACTTTTCAAA CTAGATACCC    60

```

# EP 0 841 394 A2

CTGCACCAAC CAACGTCCAT AATGGTTGAT AGTAATGATG GTCTGCTGGG TCTATAATTG 120  
 CTATTTTCTC TTTTAAGTTN TGATTTTTC TTAATAGTCT TGATGCTACG GTAACACCTG 180  
 5 CTGTACCGCC ACCAATAATA ACAATNTGAT AATGCTTATT CATTTNTTCC ACATCCTTAA 240  
 ATTATTTTAA TGATTCTGGA AAATCTTGAT AGCCTTCTCT AATATTCACC ACATTTTCAA 300  
 AACCTNTGCT TTCCAATATA CCCACTGCAA TTGAACTTCT AACACCTGAC T 351

## (2) INFORMATION FOR SEQ ID NO:20:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 324 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GAATTCGGTG TTAAGAAACA AGATTAAAG AGCAATACGT GAAAATTTC AAGTACATAA 60  
 GTCGNATATA TTGGNCAAAG ATATTATTGN TATAGNAAGA CAGGCAGCTA AAGATATGAC 120  
 GACTTTACAA ATACAGNATA GTCTTGAGCA CGTNCCTAAA ATTGGCAAAG TTTTAAATAA 180  
 AAAGATTAAG TAAGGATAGG GTAGGGAATG AAAACATTAA CCCCTCAACG CATCCCGAAG 240  
 30 TCTTACCTCA GACAAACGTT AGACTGACCC TAGGGTTAAG ACTTGGCCCN AGGNTATNN 300  
 TAACTTACTT TAAAATGTTT TCAC 324

## (2) INFORMATION FOR SEQ ID NO:21:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 484 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GAATTCCTAT TATCGAATTG CATGCTACCC CTCCTAAAAN CCAAATGCTC TTAATTTATC 60  
 50 TACAGATAAT TGGTCTTTAT CTTTATTTAA AATATTTTCA ACATATTTAA ACAAAACGTC 120  
 TGTTCCTAAA TGTACTTTAT CTCCTAATTT TGTGGATGAT AAAATCGTTG AACGCCTCGT 180  
 TTCTGGAATA AGATGAATGT CAAAACGTGT ATCATGCTTA TCAAATACCG CTAGACTTAC 240  
 ACCATCCACA GTAATAGACC TTGCTTAACT AACTGATTAT TAATATGTTG GCTACATTGA 300  
 55 ATCGNAATAA TTGTGGCATT GGTGTTTCAT TTATTTGTGA ANTGTCTAG TCACTACATG 360  
 ACCGAGGAAA AATGTCCACC AAACCTACCG NTACACTCAT GGAAGNTCTA AATTACTTCT 420

EP 0 841 394 A2

GATTGGCGCT TACANCTGCT AATAGGTTTA TTTCAAGGCT TTATTACTTG ACAGGAAAGA 480  
TGCT 484

5

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

10

(A) LENGTH: 481 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

20

GAATTCCTTA CTGTATGATA AAGGTATTAC CAAACAATAC TTAAGGGGGA TTATATACTG	60
TGGNTCAATC ATTATATGAG TTTTATAGAG AAAATATAAA TTATCTAAAA GAAAATGGTN	120
TGTATAATGA AATAGATACA ATTGAAGGTG CAAACGGACC AGAAATCAAA ATCAATGGGA	180
AATCATACAT TAACTTATCT TCAATAATT ATTTAGGACT AGCAACAAAT GAAGATNTGA	240
GATCAGCTGC AAAAGCAGCT ATTGATACAC ATGGTGTAGG TGCAGGCGCT GNTCGCACAA	300
NCAATGGTAC ATTAGATTTA CACGACGAAT TAGAAGAAAC ACTAGCAAAA TTTANNGGAA	360
CAGAAGCTGC AATAGCTTAT CAAGNCAGGA TTTANTTGGN NATATGGCTG CTATNTCAGC	420
TTGTCAANGA ATNAAAATGA TGCTATTTTN TCAGATGAGC TTAATCATGC ATCAATTATT	480
G	481

30

(2) INFORMATION FOR SEQ ID NO:23:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 459 base pairs

(B) TYPE: nucleic acid

40

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GAATTCCTTG TTAGGTGGTA ACGTTACAGA AACACATACC TTTAAATTGC CAGAAGATGC	60
TGGAGAGCGC CAGATGTTCA TTATTGATAA AAAAAGACAG ACGCCGAAAA AGTACCCAAG	120
AAAACCAGGG ACGCCTAATA AGACTCCTTT ACTTGAAAAA TAATGCATAA TCCTTTACAA	180
CTAACATAAA AGGAGCGAAT GGATAATGAA AAAACCTTTT TCAAAATTAT TTGGTTTGAA	240
AAACAAAGAT GACATCATTT GACATATTGA AGAAGGATCG CAATAGTAAT GTTGAATCCA	300
TTCAAATTGA ACGTATCGTT CCCAACCCGT TATCAACCCA AGACAGGTGT TTGAACCCAA	360
ATAAAATTAA AGAACTTGGT GAATCAAATA CAATGAACAT GGGTTTACTA CAAACCTATT	420

55



GTTGGTAAGA CCGATGAGAA GATATGTTTG AAATTATTG

459

## (2) INFORMATION FOR SEQ ID NO:24:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 440 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GAATCCCNAA	AAGATACNAT	NCNATTTAAA	TGGATATCAT	TCCAAGAAGC	TGTAGATAAG	60
AATCGGGAAC	NAATGTCCTCT	GTTTAATGAA	ATTCTTAAGA	GTCATCAAAA	GACTCCANTT	120
TTTGATACAT	CTAGCTTGCG	ACATACTTCG	AAAAAATTAT	CAACAATTTT	AAAAGGGGAT	180
TTGTAAAATG	AATATTGAAA	AAATCGAACA	AATAATCAAA	TTAGTGAAGG	AAAATGATGT	240
TAAGAAATTT	AAATATAAAA	ATTTTGANGA	TGAAATAGAA	ATTGACTTCA	CTGACTCGAA	300
TCATTTGGCT	GCACACAGGT	AATCAATCAA	NTCAANGTAT	GAACAATAAT	GATTTGACAG	360
CTTCAANAGC	GAATGATAAC	TCCGATGTNT	CGACAAATTG	ATTATCATGA	CATTAANTCA	420
CCANTGGNAG	GTACATTCTT					440

## (2) INFORMATION FOR SEQ ID NO:25:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 358 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CCCCCGGGCT	GCCACGGTAA	CTGATGCCTC	CGCGGTACTG	GCGACAGATT	ACAAAATCTC	60
GTTTCGATAAT	AATCACTGGC	AGGTCACCCG	CCTTGCCAGC	AATACCACTT	TTACGGTGAC	120
GCCGGATGCC	AACGGTAAAG	TGGCATTGTA	TGGTCTGGAG	TTGACGTTTA	CAGGAACGCC	180
TGCCGTTAAC	GACAGCTTCA	CGCTGAAACC	AGTAAGTGAC	GCCATCGTCA	ACATGGATGT	240
ATTAATCACC	GACGAAGCGA	AAATAGCGAT	GGCGAGCGAA	GAAGATGCGG	GTGATAGCGA	300
TAACCGCAAC	GGTCAGGCC	TGCTGGATCT	TGCAAAGCAA	CAGTAAAACG	GNNGGCGG	358

## (2) INFORMATION FOR SEQ ID NO:26:

# EP 0 841 394 A2

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 480 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GAATTCCTAC	CAGACGTCGC	AATGGTCTAT	AATTTGCTC	TAGATAATCT	TGGAATAATA	60
TCTTTTATTG	TTAAATATAC	TTCTGGTTGT	CCAAACATTG	CTGTTGAAAT	AGCAAAGTAA	120
CTTTCTAAGC	GCCCCATTCT	CGTTATTTTA	TTAATAGCGA	TACCTACATA	TTTGATAATA	180
AATGGTAATA	CCTTAATATA	ATTAAAAATG	CCTATTAAATA	CAGAAATGAA	CACTAATGGC	240
AGTAACACAT	TTAAAAAGAA	CGTAAAGCCA	TTTNTATTTT	GTATATCTCC	AAAAACAAAG	300
TTTATACCTG	CTTTACTAAC	ATTTATTAAC	CCTTCCAAAA	AATGAGCCTA	ATGCAGGTAA	360
GATTGGTTAA	ACCAATCGTG	GTGTTTCATCC	ATTAAATTAC	AACAATCCAA	CAACTTTGGA	420
TTAAATTATC	AATTATTAAC	NGCCNTTTTT	GGAAGTCTAT	GTTTTTACGG	ATCCCAAAAA	480

## (2) INFORMATION FOR SEQ ID NO:27:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 545 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GAATTCCTCT	TTCACTTCTA	TGATTTTTTT	TAACTTCGAA	ACGTCAAGTT	GTGCATCAGG	60
TTCTGCTGTT	ACTTCCATTA	CATAACCTTG	AATCGTTCTT	GGTCCAAAAG	GTACAATTAC	120
ACGCACACCA	GGTTGGATGA	CAGATTCGAG	TTGTTGCGGA	ATTATATAAT	CAAATTTATA	180
GTCAACGCTC	TTGACGCGA	CATCGACTAT	GACTTTCGCT	ATCATTATTG	CCACCTAGTT	240
TCTAGTTCAT	CTAAAATTG	TGCAGCTAAT	ACTACTTTTT	TTCTTTTCTT	GATATTTACC	300
TTTTTCATTAT	TTCTAAAATG	CATTGTCAAT	TCATTATCAT	CAGAACTAAA	TCCGATAGAC	360
ATATCCCCAA	CATTATTTGA	AATAATCACA	TCTGCATTTN	TCTTGCGTAA	TTTTTGTTGT	420
GCATAATTTT	CAATATCTTC	AGCCTCTGCT	GAAAGGCTAT	TAATACTGNG	ATGTTTATGT	480
CACCTAAATA	TTAAGAAATG	CTTCAGACGT	TTAAAAGATA	CTGACAATCA	CCACCTGGTT	540
TTTCA						545

## (2) INFORMATION FOR SEQ ID NO:28:

EP 0 841 394 A2

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 429 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GAATTCCTNA	CCATCCATAG	CTGTGNACAC	AACATGATCT	TTATCAATCA	TCATAATGTT	60
GCAAAGGGAT	TTCTTTNATC	AAACTCAGGT	GCATATGCTC	GNACTTCTTC	CATAGTTANG	120
GGCAATTTTCG	TTGGGGTTAA	AGGGTGTGTT	GGATGCCCT	CTGGNNCCAA	GCTCTCAGAA	180
TAAGTTAACT	CATCTGTCTT	ATCACTATCT	TTTACATGTT	GTANCCATGA	AAAAAAGTTA	240
ATATCTTGTG	GTAATCTNGA	AAATTTAAAA	CTGTGACTTA	TTAATGTGTG	TCTGTGAGAC	300
ATGTGNTTAT	ATGTCACATC	TAAACTATCT	CGACTGTGTA	TTANCTCCTT	CGNNGTAGTC	360
GCTCTGANNNA	NTAGGGATAT	CAANGGTCAT	GACTCAACAC	ATCTAACAAC	NCCTTCTTAA	420
NGGTTATTA						429

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 434 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GAATTCCTAA	TGCANAAAAAN	TTCCCCGAGT	CACAGTACAA	TTTGCCGTGC	TTCGGNGGAT	60
TACTTTGTTT	ATTTTATTG	TGTANCATTA	ATAATGAGTT	GTCTTAACGA	TTGTCGCTTT	120
ATAACTACAC	GTGCTGTGAC	ATCTTTTAAG	AAAAACACAG	AAGGCTTGTG	CATGTGATTG	180
TAATTGGATG	NCATAGAGTA	ATGATATGCA	CCAGNTGATA	ATATAGCAAG	ATAGTCTCCA	240
CGTGTGACTG	ATGAAGGTAA	TTTAGCGTCT	TTTAATAATG	ATATCACCAG	ACTCACATAA	300
TTTCCAGCT	ATAGNTACAC	TGACATCTGC	TTCTACATCT	CTTATTTAAC	AAAGCAATGC	360
TTGATACTGT	GCGGTCATAA	AGGTGAAGGT	TCTTGATATT	GAATCACTCA	TNCCGCATCG	420
AGTTGAACAT	ATTT					434

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 535 base pairs

EP 0 841 394 A2

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GAATTCCTAC	AAGGATTCAA	AATAAAATAC	AAGTGTTTGT	CGAATATTTG	TTCGTATAAA	60
ATCTTGAAAT	CAAACGCAA	AAACGAACAA	ACGTGCTTAT	ATTTTATCG	ATAATACAAT	120
TANNTTGTG	TTCTAAGCCA	ATTTAATAAT	CTTATGAGCA	CATNATTTTG	ACTTCTGTTC	180
TTTAATAAAT	TTCTAGATT	AGACATTTTA	AATTGCTCAG	TTTTAAATTC	AAAGTTATTA	240
TCATAGCCTA	TATGGACTAC	ACCATCATAC	TCTAATAGAA	TAATTGCTGA	TGACACATTA	300
AACAATTTAT	TAACAAATTG	CACTGCATCT	TTTAATTGCT	GCTCAATAGG	TTTATTAATA	360
TCAACAAAGT	TTTCATTGGT	GATTAATCAT	ACCCTTTAGA	ACATCGGNTT	AAATCCNCTT	420
CTTNCAATCG	ATGATATTAA	GCACCATTAG	TAATACCATC	ATTTTACAA	AAGGGTCANG	480
AATTTTCTTN	ATTACAGCTT	GGTCAATTAT	GTGTCANCTT	GAACCATARR	DCTNS	535

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 376 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GAATTCCTGC	ATTTCCAGGA	GGATANGAAA	TAAGAATTGA	GTTATTAATC	AGTGGTTTTA	60
TTTTTATATT	GGCTTCCCTT	AGCGATTGTG	GTTGATGGTT	ATTTAGCTAG	AAAATGGAAT	120
TTAGTTACAA	ATATGGGGAA	ATTTGTGGGN	TCCATTAGCG	GATAAATTAT	TAGNTGCAAG	180
TGCTTTAATT	GTNCTTGTGC	AACTAGGACT	AACAAATTCT	GTAGTAGCAA	TCATTATTAT	240
TGNCAGAGAA	TNTGCCGTAA	CTGGTTTACG	TTTACTACAA	ATTGAACAAG	GATNTGGTAA	300
GTGCAGCTGG	NCAATTAGGT	AAAATTAACA	CAGCAGTTAC	TATGGGTAGC	AAATTACTTG	360
GGTGGGTATT	AGGGTG					376

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 518 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single

# EP 0 841 394 A2

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GAATTCTCAG	TTCATGTTGT	GGNGACACTT	TAAACGGNCT	GTGCCAGTAG	CGACCGAGTC	60
ATTTCAAGAA	TGACCATTTC	ACATTTATAT	TATAACACTT	GTCGNGCGTA	ACTGTATAGN	120
TTTTCAGNNG	TGTTTAAAGT	TAAGTTATCT	ACTTCGCGCT	TTCCCTGGCT	TAATTGTGAA	180
ATTACATATT	GCGCTACGCC	AGNTTGGTTG	GGAATTTGGT	AACCTGTTAT	ATCACTTNTG	240
GTCAANNCAA	TTATTTTAA	TTTATAATCA	CTCATATTAT	CTACGTCCAC	TCNTTTTATC	300
TAAACAATAA	AAATGTGTTT	NTCTCCCCNT	AAATAATAAC	AATGGCAGGC	TTAATAAAAA	360
CAATATTAAA	TACAAATTGC	TCTTGTCAAA	ATTGAAAACC	CCCACATAAT	ATTATTTTAT	420
TTAAGNGTAA	GAGGGGGGCA	ATCAGGGCTC	CAGCATAATG	TTAATCTTTG	TCGGGTGGGG	480
TTCCNGTCTT	GGGTNGCCCN	AGATGCCAC	TCTCTAGG			518

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 423 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GAATTCCACT	AGAAGTTNCA	ACATTACGCA	TTTACATTT	CGTAAGTGGA	GACGGTGGCA	60
AGCTAAAGAT	TCATTATGAA	TTATATCAAG	ATAATGAAAA	AATGGGTTCT	TATCAATATG	120
AAATTAAC TA	TNAGGAGATA	GGCGAATGAA	TATTATTGAT	CAAGTGAAAC	AAACATNAGT	180
AGAAGAAATT	GNAGCAAGTA	TTAACANAGC	AGGATTAGNA	GNTGAGATTC	CTGATTTTAA	240
AANTGAAGTT	CCNNNGNNT	NCAAAAAATG	GAGATTATGC	TACTAATATT	GCGATGGTAC	300
TGACTAAGAT	TGGAAAGNNG	TNCTCCNCNC	TGNNNATTGC	TCANGCGGAA	NGNTTGNNNA	360
CCTTAGCTCC	NGGACACCNG	CACATGNAAC	ACAACTNGAC	ATTGCCTGGC	CCCACGGCTT	420
CAA						423

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 943 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

AATGACAAAA TATTAATCAA TTGGCGTATT ATACAAGAGT AATAGACATA TAAAAGTCGA 60  
 ATGTAAC TAC GTGAGTACTG ATTTTATTCT TTGTAAATTA CAAGCATTTTC ATATCATAAA 120  
 GTTTGAAAAG AGGTATATTG AAGTGGAGAA AAACGAATAT ACAGCTAAAT ATAATGAATA 180  
 TAGTCAATTA TTAGACGCTA CTTACTCGCA AGCTGTAGCA TACCTTTTAA ATAAATATGG 240  
 CGCTGTAACC GATGATTATT ATAAGGAAAA ATCATAACACG CGATTTTAA ATGGAGAAAT 300  
 CAAAAGTATT TCAAAAGGAA AATACACTAG AGCTAGCGAA GGATTATATT GTCATCACAT 360  
 AAGCGAAGAC AAATTCCTAAA ACTTATCTTG ACCTAAGATT CATTTCCGAA TTAAAGTACT 420  
 CATACAACGT CAAAAGAAAG AAACTTAGT GTACTGTGAT CTTATCGAGC ATTTANTTTT 480  
 ACATGCAATT ATTACAAAAG AATCCAATGG CAATTGGGGT GTGGGCTGGA TTATGTCAAA 540  
 TGATCAAACC AACAGTCATT GATTGGGACA TTAGCGAATA TACTCCAAA CCAGCATGGA 600  
 TGCAAGCCAC CAAAGCACGT GCCTATGTGC CTAGAATATT AGTAGAGAAA CTACTCATTA 660  
 NNATTGACGA TATGTTAGAA GGAATAGAAA TATATGATNT CCNTGAGTCT AGATAAATGA 720  
 TTAANGGAGA TATATTTCTG GNTCGNTGAG ATTGTGNTAT AGATAGNNTA ACCCNNTAAT 780  
 TAAATATTAT TCTATNAAAT AGACCACGCA TNCCTATCTA TAAACGGACA ACGTTTATAA 840  
 ATGAGTTTGC ATGGTCTTGA ATTGTATTAA ATNTCTTNTT GGGTTTAAAT AGNTCACTTA 900  
 GNGGGTATGT CAACATTCTT TNGACAGCTT CACGTGTATT TTT 943

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1119 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CAGATATACC TGCCTAATC AAAAGTTTCAC CAAAATTTAA AGCTGTGTGA TATATCACAA 60  
 ATCCACAAAA TCCTAAAATA AAAATAACAG GGATATCTCT TAGTTCAGGG GTAGGCAATT 120  
 TCTTTATAAT TACGAACGGC AAGAGAATTA TTGTTGCTAA AATTAAACGA AATGCCGACA 180  
 ATGATTCTGC ACTAAAATCA TTAAACGCAA TCTTTATCAT TGGAAATGCA GATCCCCACA 240  
 ATATGATAGT AAATAAATAT GATAGAAAAG TAGTGTCTCG AAGTTTATTC ATTAATATCA 300  
 TCACTCCTTT AATTATGTGT TTCTATATTA AAAAATATGA TTAAAAATGA GTACAACCAA 360  
 TTGGGAATGG ATTTACCTAT CCAATTTTAA AAGGGAGGGA GAAGATGGCT AAATATTAAG 420  
 ATATTGCTAG TGACATAAGA GATAAAATAA TCACAGGGGG ATTGGGTTT ATGGAATGAA 480  
 GATACCTCCA CAGAGGCAGG TGGCGATACA AATACAACGT NAATAGAGTG ACGATTATTA 540  
 AAAGTATTGG AGTTATTAGA GGCTGAAGGA TTTACTATAC TAAAGTGGGG AGTGGGNACA 600

# EP 0 841 394 A2

	TATGTTAATG	ACTATTTGAA	TGAAGCACAT	ATTACAAATA	AGTGGTCTGA	AATGATGTTA	660
	TGGTCCTCTC	AACAAAGAAG	TCAGTATACG	GTGCAATTAA	TTAATAAAAT	TGAGACAGAT	720
5	GATTCGTATA	TACATATAAG	TAAAGGTGAA	TTGGGTATAT	CGTTAATGCC	ACATATTCAA	780
	TTGAAAAAAG	CCATGTCTAA	TACAGCCAGT	CATATTGAAG	ACTTATCTTG	TGGGTTATAA	840
	TAATGGCTAT	GGTTATATCA	AGGTTAAGAG	ATATTATCCG	GNTGAACGAA	TTGTCAAAGC	900
	AAGGTATTAA	ATTGTAGGTA	GAGGAAAATT	GTAATGGATC	AACTCCAGGC	GCTTTTACAA	960
10	TGCCANACCA	ACTNTTTATC	TTATTGGGGT	TTNTAGGGTC	AAAGATGCCA	TAATAATTTC	1020
	GGATACAACC	ATCATATTTT	CCACTCTACA	AANTGTTTGN	GGAACAAATG	GAATTTTAGA	1080
	CATATTGATG	TCCCCNTATT	ATCAAATTTA	ATGAAATTA			1119

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 600 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

30	AAGAAATGTT	CTAGAGATTT	AGGTATCGAA	GAACCCAGGA	TTAAGACGAT	TAAAAGAGCN	60
	ACTATGATAT	TAGGATTATC	AACATATTTA	CTGCTGTGTG	CAAGAAGTAC	GTGCTTGGAC	120
	AATTTAAACC	AGGTATGACT	GCACCTCAAT	GTGCTGGTAT	CATTCATACT	GATTTTGANC	180
	GTGGATTTAT	CCGTGCCCGA	AGTAACAAGT	TATTGATGAC	TATGTACAAT	ATGGCGGCGA	240
35	AAGTGGCGCT	AAAGAAGCGG	GCAGACACCG	ATTAGAAGGT	AAAGAATATA	TTATGCAAGA	300
	TGGCGATATC	GTTCATTTCA	GATTTAATGT	ATAAACGATA	GAGTGAAGTT	AATTAAATAG	360
	TATATATGTA	GAAGAGGCGG	AATCAATTGT	TCGCCTCTTT	TAATTATGCG	TATAATTTAT	420
	TAAAAGAATG	GAATGATTTT	ACTCGCGTAA	ATAATATCTT	GAGTGCTGAA	AAATTGTTTG	480
40	CCTTCGCCAG	TATAAGCAGG	CTCTAAAACA	AGATTAGCCT	TTGCACAATA	AAGCCATTCA	540
	GGATGAATGC	CACATATTAAG	TATCTCTTGG	AATTCGATAT	CAAGCTTATC	GATACCGTCG	600

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 479 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

# EP 0 841 394 A2

	GAATTCGNAA ATGTTGGCTT TGGCGCAATG AAGAAGCAAC CACCNTCTAA AAGCTAAGTT	60
5	AANCTTATTT GAACACATGA TNNCGANTGC ACGTACATCA NCAGAACAAG GAATGCTTGT	120
	NTCTGCGGNA ACAATGCTCA NNCANAGTAC GCTACCCACA TTTACGAAAA ATACATCACG	180
	AAGTCGTTC A TGATATGAAA TGTAGCAGAC ACGAGATTCA CCCTNTGTNT CCAGCTGACT	240
	TCGCAAGTCA NNNAAATGTG TTAACCTCTAG CGGAAATGAA AAGGACAATT CATGATATAT	300
10	TGGATTTCAG AGATGAAGAT ATTTGGGNTT GTTATTTGGG TACTTTGGCT GTGTCTCCAN	360
	GCTTAANGGA TGATGANGGT CANMNGTTAC TATCACGTNT ANGAAATGCT TACAATNNNTT	420
	ACGANTAGAA ATGACCCGGA TGGCGNGGNC NGNTAAATTG TCAATGNNTT NAAAGGTAA	479

## (2) INFORMATION FOR SEQ ID NO:38:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 733 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

30	GCGAGCTTAG TTCATTTTAA AGTAAAGAGT TTTTGTATG TAAAGTATAT TAGTAAAACA	60
	CAGATTACTT GTCTCTGTGA AATTAATACG TTTATACTAA AAGATATAGT CTTTCTAGAA	120
	AATCTGTTAA TTAATTGTGA ATTTTATAGAA AATTTATTGA ACAGCAAAAT ATGGATTGTT	180
	ATAATTTAAG TTAAACAAAT TCTTATACAA TATTATTAGG AGGCAATCAC CATGTCACAA	240
35	GCAGTTAAAG TTGAACGACG AGAAACATTA AAACAAAAAC CAAATACATC TCAACTAGGT	300
	TGTGGGTAAA TATTTTACTG ATTATATGTT GAGTTATGAT TATGATGCAG ATAAAGGATG	360
	GCATGATTTG AAGATAGTAC CTTATGGGCC TATTGAAATT TCACCTGCTG CACAAGGTGT	420
	TCAATTATGG TCAATCGGTA TTCGAAGGAT TAAAAGCATA TAAAAGAGAT GGGGAAGGTT	480
40	GCACTTTNCC CCNNCCCGAA GAAAATTTTA AGCGTCTTAA TAACTCGTTA GCACGATTAG	540
	AAATGCCTCA AGTAGACGAT GCAGAATTGT TAGAGGGGCT AAAACANTTA GTTGATATAG	600
	AAAGAGATNG GNTTCCTGAA GGGGAAGGTC AATCATTATA TTTTCGNCCA TTTGGTCTNG	660
	CAACAGAAGG GGCACCTNGC GTTGGCGCAT CACATCAGTA TNAAATTTAT TAAATTATNT	720
45	TATCTTCCTT CAG	733

## (2) INFORMATION FOR SEQ ID NO:39:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1263 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear



# EP 0 841 394 A2

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

```

5      ACTTGTGCAC TGCCAATAAT AATTAATCCT TTCATATTTG ACCCCTCTTT TTTTATTAA      60
      ACTTGTGTGT CACGTAATAA TTTCATAATA GCTTTTCCAA CACCACTATT TTCATTGTA      120
      TCTGTGACAT ATTTGCTAT TTTTAACT TCTTCTGCAC CATTTTCCAT CGCAACTGGA      180
10     TAGCCAACCTT TCTCTAACAT TGATAAGTCA TTTAAATTGT CACCTATTGG CATGACATCT      240
      TTCATTTCAA TCCCTAATCT TTCGGCAATT GTTCTAGCG CAATACCTTT NGGTGCATCT      300
      GAATGCGTTA TTTCTATATT TCCTCTCGAA GATGATGATA TAGCTAAATT CGGAGATTCA      360
      GCTAAAATTT TACTAGCTTN GGCAATTTTT TCTTAAATTT CCATCAAATG CTAATATTTT      420
15     CATAATTAAT TCAACAGGTA TGTTTTCAAT AGGATCATAA TTATCAACAA CTTTCAACGG      480
      TNCCAATATC TATGCGTCTT TGNATACCAT CTTNAATACG CTCAACGTNT GGATGTTGAC      540
      CTGCACGCTC AGCAATATCT ATGTAAATGT CTAAATCTCT TGGGGGATCT TCCAGTATAA      600
      ATCGCACGAC TCGGNTATNC TGATAATAA GTACCTGNAT CTTTAAAAC ATNTGTAATG      660
20     NGGTGTACTA ACGATTATTT ANGGNGTGAA GTGCTCATT CATTGAAAGT TTCATCACGT      720
      TCTTCAGNAC CATTGAGACA AATATATGGT ACTGNTAAAT CTGTGTCAGC AANTGGTGCT      780
      TGTGCTTCAA AAAATGCTCG ACCTGGCACG AGCGATAACA ACCGTTATCC CTTGTTCTTG      840
      AGCGTATTNA NTCGCATCAA TATTAGGTTG AGAAATMCA TMTGCTGCAT TAAGGTAGCG      900
25     NGCCATCCAT ATCAGNGGNT ATTAGGTTTN TCANTATGTT NCCNCGCCCN CGNCTTTCGT      960
      AAATTTAAAA TCTTGTTCTT AAATAAGAAT ATATACTCAG CGCACATACT TTTCTATTAA      1020
      CATTTATATT GTCATTAATT TATCATATAA TGTAATTCTA ACAAATTTTA ATTAGTATGT      1080
      ACTATCGTCT TAATGGTGG ATTTCTTATT GGCTCTTAAG TTTTAAAAA ATGTTGGTTA      1140
30     ATAATGGTGC TACATGCTTC TTTAAGGTA CAACCTTTAT CAACAAATG GACGGATGAT      1200
      TAAAAAATTAG GTTNGGTTGG AAATAAATTC AATTAACCTT GCTACTACAA ACCACRRDC      1260
      TNS

```

35 (2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

```

40     (A) LENGTH: 594 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

```

45 (ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

```

50     ATTATGCCTA TTTGACACAT TATTGAAATT TTAAATATAC CCCTTATACG TTACGGCTTT      60
      GTACTGTGTT TAGGTATATA AAGCTTTGAC ATATCGATAT TCTCCACTC TAACAGCTTA      120
      ATTTTATTAT TAATCGTTCC ACCGAATCCT GTTAAGCTAC CCGTTTTACC GACAACACGA      180
55     TGACATGGCA CGATAATTGA TAATGGATTA CTTCCGACTG CACCTCCAAC CGCTTGGGCT      240
      GACATTTGTG GCTTATCAAG CACCTTACCT ACTTTGTGG CAATAGCACC ATACGTTGTT      300

```

EP 0 841 394 A2

AGAGTCCCAT	AAGGAACCTG	TCTTAATTCA	TTCCAAACAC	ACTGTTGAAA	ATGACTACCT	360
GTTGGCTTTA	AAGGTATTGT	GATTTCAGGA	TTGTCAACCT	TTAAAATACG	CGTCTTAACC	420
ACTGTGTCGN	CTCTCTAAAT	ATCGCTAAAG	ACGTATTTTC	TTCCCTAGTA	CCAACACCTT	480
GTTGATTTTC	AAACAAACAG	CGGTCAGATT	ACCCATCACT	CAAAAGTCCA	ATCGTCTACA	540
GCGATCATAG	TACTCTATCT	CCATAAATCC	CCCTTTTCTT	ATCATACGCC	TATG	594

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 511 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

AATGGGAAGT	TCAATACCCA	TAATGNCAAG	CACCCTTGCA	TTTAATTGCA	TTTTCCNACC	60
GTTATTACNG	GCCCCATCTC	CAACAATATG	TAAAGTTAAT	ATCTGCATAA	TCTTTTTTTA	120
AAGACTTCTT	TTGCTGTTAA	GAAATCAGCA	AGANCTTGAG	TAGGATGATC	TTCATCAGTT	180
AAACCATTCC	ACACTGGTAC	GCCTGAATAT	TCAGCTAAAG	TTTCTACTGT	TCTTTGTGTN	240
AAACCCGCGG	TATTCAATGC	CATCATACAT	TCCACCAAGC	ACACGTGCAG	TATCTTTAGT	300
TGTTTCTTTT	TTACCCATTT	GTGATCCAGT	TGGGCCTAAA	TAAGTTACAT	TTGCACCTTG	360
ATCATGCGCT	GCAACTTCAA	ATGCACATCG	CGTTCTTGTA	GAATCTTTT	CAAATAACAG	420
TGCAATATTT	TTATTTTFTA	ACATAGGCTT	TTCAGTGCCA	ATATATTTAG	CACGTTTTAA	480
ATCCTCGGAG	AGNGTTAATA	ATNNGNAATT	C			511

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 546 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

CTGCCTTTTC	TCTATCTGAT	AAACCACGCT	TTATTAAAT	TAAATCTCCT	AACCATACCC	60
CGTCTAAATT	ATCTGGAAGT	ACATCAGCCT	TCCTTACTTC	AATATAATCA	TGTTGTATTA	120
AAGTTTCTTC	ATATAATCCC	ATCTGATACA	TCCTTTACTT	ACGTTTACTT	CTTATATAAT	180
CTGCATAATC	TAAAACTCTT	TGGCATT CAT	CATCTGTCAA	TTCTCCTTCA	AGGTGAGCTG	240

# EP 0 841 394 A2

CACGATGTTG	TACTTCGNTT	TCTGGTTGTC	TATTTNNTAA	TAGTAAATAT	TCTGGGGTAA	300
CTTTCAATGC	ATTGGCAATT	TCAGCTATAT	CCTCCATAGG	TATTTNTCTG	CTACCGNNTN	360
CGTATCGGGA	TAAGGTAGAT	TTATTGACAC	CTATCTTAGG	TGGAAAATCA	GGTAAATTCA	420
CATTATTTNC	TTTTCGGAGG	TTGGGTGATT	AATTTANCTA	TTCCCGCTGA	AGTNCTCATA	480
TTCAAATTTA	CTCCCGTTTA	TTTATACAGT	ATAATAACAC	TTCCCATATT	AGGAAACACT	540
AGCATT						546

## (2) INFORMATION FOR SEQ ID NO:43:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 572 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

TTATATTATG	TTCTGTTGGA	CCCAGTAAAA	TGTTGCAATG	TATGATAAAA	AGCCCTTGGT	60
TCAATTAATC	GAAAATCGAC	ACCAATCGTT	TCATCATCAT	ACGCTAATTC	TAATGAAACT	120
TGTGTGTTCC	CCACTGCAAC	TTCTTTTACT	TTATTGGGAA	TATTTAATAA	TTGCTGCTGC	180
ACTGCTTG TG	GGTTATCGGT	ACTTATTATG	AAATCTAAAT	CTTTGCTCAT	TTCTTTAAAA	240
CGACGGAAGC	TTCTGTCAGA	TGAATATTGA	TCGATATAAT	TTAATGTATC	TATATAATCA	300
ATGATTTCTT	GATTAAGTCT	TCTCATTGGA	TCAATTGGAT	ATCTATCTTT	CTTAGCACAG	360
NGGGGTTTCA	CAGCTTCTAA	TATGGTGTGG	TCCGNTTCTT	TTAGCAAATC	CGCTTAATTC	420
ACTTAAC TTN	TCCCAATNTC	ACAANCAACT	TGGAAGTGAC	GCTTNTATCA	ACAAATATTC	480
CAACTCTTTA	TATAGCTTAG	CAATTTCTTG	GTTC CAAAGTC	TGGAANTTTC	AAAAGTGGAT	540
TAAGACCTCC	CGGACTTCTT	CCCGGTAATG	GT			572

## (2) INFORMATION FOR SEQ ID NO:44:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1115 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GGATGCACTC	TGTTAGCATT	ATCTTTAAAC	ATTAAATATG	GATAACCTGA	TTGTAATTGT	60
GTNTGCGCAA	TTAAATTCAA	CATTTACGCG	GCATTCTTTT	TCTTTTCTC	AACATTTGGG	120

EP 0 841 394 A2

5 TTTGCAACCA TGTCATCATA ATATTTTCT AAATCGATAT CGTCTAATGT CACACCATAT 180  
 TCTTCTTTAA CTGTATGAGG TGCAAACATA TAAAAGTCCT TACCTTCTTT AGCTAAATCG 240  
 AAGAATTTAG ATGGAACAAT TAAACCAGTT GATATTGTAG ATAAACGTAA ATCTTCATCC 300  
 GCATTTACTT NNTTAGTATC TAAAAATCT TCTACATCAT AATGGAAGAT ATTTAAGGTA 360  
 CACAGCACCA GCACCAGGGC GGTGGACAA GTTGGATCTG CATAGCTAAA GCCACCTTCA 420  
 AGTGACTTAG CAATAGGTAA AACGCCTTNC GCTACGCCTT TAATTCCTTT AATTGCTTCA 480  
 10 CCACGTGCAC GCAATTTAGA TAAGTTAATT GCAACGCCGC CCCCAATTTT ACTTAATTGG 540  
 TNGNGCAAGG NGGANTCAAT AAAGTTAANT TGAATTTAAG NTGNCAACCC ACTTCTTNNA 600  
 TNNGAAACAT GACACTAGCT CACCACGACG CGCACGGNCT GCGTTTAAAA ATGTTGGTGT 660  
 CGCTGGNTGA TATCTTTGTT CAACCATAGC AGAAATAAAT TGTTTAGCTT GTGCTTTATT 720  
 15 ACCATNTGCT AGGTATAAAG CAACAATGGC AACGTGTTGA TTATAGTCTT CTAAGTATTG 780  
 ACTTTTATCA TTTGTTTNC AAGCGTAATC TTTGAAAAAT TTACTAGCTG ACATATAACT 840  
 TGCAAAATTA AACGAGATTG ATNTTGGATA ATCAGTGATT TCAATTAGAT CCNCTTCACT 900  
 ATAAATATCA AACACATTGA AATAAAAAATC GGTGCTTCT AAATAACGTA AAACGCTCGA 960  
 20 TTTCATATT CGAAGAAGAT NGCGTGGGCT TTTTACTTCT TTCTTAAATT AAGCTTACNT 1020  
 AAAGGTTTCT TGGGGCTTTT TTCTTAAACT TAAAGGAATC CCATCTTCCT CTTCCTCGGT 1080  
 NGGGGACCCC AAATATTTTA CTCAAATATT GATGG 1115

25 (2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 1123 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

40 TTCCATGGAG GCATGCGCAC AATCTGCATT GTAATTCATA CCAAAATGTT CACCGACTCC 60  
 CAACCGAATC CACCGAAAAT GCCAACATCT AGAATCGGTT GCACATCATG TTCAACACAT 120  
 TCTTCGTGCA ATAAATTAGA GAGTTCATTA TTGATGCCAT AACCGATGCT AATTGTATCN 180  
 CCATAAGTTA AAAACTGAGC AGCACGTCGG AGAATCAATT TCCGACTANT AAAAGGTAAT 240  
 45 GCGGGTTCAG GTATTCATC AATTCGTTCT TCTCAAGACA AGGCTGGTAA ATAATGACTC 300  
 TGAATTACTT GGCGGTGATT CTTTTCATCT TCTGTGACGA ATACATAATC GACAAGATTT 360  
 CCTGGGATAA CAACTTCATT CGGTTTTAGG TGATAGTCGT CAACTAAAGC TTAACTTGN 420  
 ACAATAACTT TCCCATGATT GGCTTTCGCG TTAAATGCGA CATGATAACA CTCGCTCAAG 480  
 50 TACGCTTCTT GAGNTAAATA AATGTTACCT TGTGTATCTG CGTATGTTCC TCTCAGTAGT 540  
 GCCACATCAA CGCTAGGGAA TGTGTAATGT AAGTATGTTT CATCGTTGAT GGTACTAAT 600  
 GAAACTAAAT CATCAGTTGT TCGTGTATTT ACTTTACCGC CACCGTATCT AGGATCAACA 660  
 GCTGTGTTTA ATCCGATTTT AGTAATAACT CCAGGTAATA ATTGATTACT CTGACGATAA 720  
 55 TGAGTTGCAA TGATACCTTG TGGTAAAAA TAAGCTTCAA TGTCATTATT TTTCATTGCT 780  
 TGTGCCGTTT TGAAGAAGC CGTTAAATG CTCATAATGA CACGTTTAAT CATGCGACGT 840

# EP 0 841 394 A2

TCTATAAAAT CATCTAAATC CGGTGCGGCA CCTAAACTAT GGATATCATT CGCTAATATA 900  
AACGTTAAAT CTTTGGGCGT ATGATATGTG TCATGTTGCG CTAACACAGC ACGTAGAACT 960  
TCGGCGGGTA AGTTGGCTAC AGCTAATGCT GGTAACCAA TCACATCACC ATCTTTAATG 1020  
ATATGTTGTA AGTCGTGCCA TGTGATTTGT TTCAAGCAAG TCACCTCCAT CACATTTGAT 1080  
AAAATATAGC GTTTTACAC TTTGTGTAAC CCCTTACAAG AAG 1123

## (2) INFORMATION FOR SEQ ID NO:46:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 960 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

CTCCATTTGC ATTTGGGTCT TTAAACTAA ACACATACTT ATAGTTGCCA TTATGTTCTT 60  
CAATAGAATA ATCATTATAC ACTTATTAT TACTACCAA TTTATTTGCT TCATTATTAG 120  
CCGCATTTAA AGCTGTTTGG AAATTTGGCA ATTGCTGTAA AGCTTGAGTT TTATTTCCAT 180  
TAAACGGATA AATTTGACGT GCAACCGGCG CGGCATTTTG TCCATAATAT GGTGCAACGT 240  
AACTTGATTN TTGATTATTA CTTGATTGGG TTATTATTTG TTTGGGTNTT GGGTCATTGT 300  
TTGTTGCATT TGAATTAGA TTGTGCTGG TTATCGTTTG CACTATTATC TTTATTATCT 360  
TTGNTTACGT CTTNACTATC ATCTTNATTG TCTTCTTTAT CTNNAGATGA ATCANTTGNT 420  
TTNTTATCTT GTTGNTCAGT TCTCGCTTTA TCATCTTTTT CTNNATTACT GTCTTTTTTG 480  
TGGNGCACTA TCTTGACCAC ATGCAGCTNA AAATAATGAT AATGCTAGTG GACCCTGTAN 540  
CTAATCTTTT CATAATATC TCCTCCTATA ATTCGATATT CATTGAATAA GCTTGAAATA 600  
CATATCTNCC ATGTGTATCA TTTTCATGGT TGTTACCANA GACGTCAACA TTAATATGAT 660  
TTGANTTATC ACAATACCCA TAACTACAAT CTAAGGTTCT TTTTATTATA CCCTAATTTT 720  
TGTTCATTAT TATTTAATTT TTGTGAATTT TATGTTTCT ATAAATTTAA TTATTTTACT 780  
TTAACAATTC ATTACGCATT TAGCATTTCA AGGCATACAC AATAGGAATT CCGATGATT 840  
CATTTTATCT TGCTGCAAAA AACAAATCATT ATTACTTCTT TTTTCCATAA TTAAATCTTG 900  
TATCCCGTTA CATCACCTGG TNTGGAAATG ATGTTCAACC CAACCACTTT TAAATCCAAG 960

## (2) INFORMATION FOR SEQ ID NO:47:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1130 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

5	TTGGCCCTTT GGAGANTCCG GATTATTGAT AAAGATATTC CTAGCGTTCG CTAAGCATAG	60
	TTATGTGATG GANAATAGTC TTGTTGAAGA GTTGTTC AAT ATTGCCTCGN CTGTCGCTCC	120
	AAGTAATGAT AAGCAGGGTG TTTTANCGAT TATCGAACAA GAAGTATTGN NCAAGTAGGT	180
	AGCACATCAT CAATGAAGCG GCTATCATGA TACCTAAATA CATGTNNAT GNAANCTTTG	240
10	NATACCGAAT CATCNAACAT TAATATATCT CCAAACAGGT TAGCCAATAA TCCCATAGTA	300
	AAATAAGCCC ACTACAATCA GTTGAATAGA TGATAAAATA AAGCAGATTT TAATCAANCC	360
	GATTTTGATA CAGATAAACA ANATANNAAN CAAAGGCACA TGGTGCGCCT GTGCTTTTAA	420
	ATATATATTA ATTTAAACGT GTTTCACATG TACCCAAGTT GTTAATNGAC AAGATAATTG	480
15	CTGCGTTTAA ANCCACCTT CCAACAAGGT TTGTACCTG CTTTCATCAGA GAAGAAACCA	540
	ATATGTGGTG TTAATAAAAT TCCTTTCATG TTCCGATTAA CTCTAATAAT GTTTTATCGT	600
	CAATGTCTTT ATTAGTCCAG TCATTTGTGA AGTATGCAGC TTCATTTTCA TAAGTATCAA	660
	TCGCAGCACC TAACAAAGTA CCATCGTTCA CTGCAGCGAN TAAATCAGGT GTATTGATGA	720
20	CTGCACCACG TGCTGCGTTA ACTAAGATTG CACCTTTTTT CACATGATCA AACATTGCTT	780
	TATCGAATAA ATGATAGCTT TCTTTGTTCG CTGGAACATG TAAAGAAATA ATATCGGCAT	840
	CTTTAATAGC TTCTTTAACA CTATCTTTAT AAGTTAAAAA GTCTAAATCT TTATTAGGAT	900
	AAGCGTCATA AGCTGTAATT GTAGCACCAA ATCCTGCATA TATTTTAGCT GTAGCAGCAC	960
25	CGATACGACC CGTACCGATA ATTGCAACAG TCATATTTTT AACTGGTTTA GACATGATTT	1020
	CTGCTTGCCA AGTAAATCA TGTGTTTGTA CACGGCGTTC AATATCTGGG AAGCGACGCA	1080
	CTAATTGTAG GCGGATAGAA ACAGAATACT CTGCAGTGTT TCAGGGGAAG	1130

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 573 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

45	GCAGTATCAA TAGTAGTTGG TTCATCATTTG CCCAACACCT TTTTCAAATT ATCATTTTCA	60
	AGTTGAAGCG CTACATTTTC TTCTACTAAT TCAACTGCAA GTGCCTTTAA TTCTGAAGTT	120
	TCCTTTGAAA GTTGATTGAC ATTCATTTCT AAACGCATTA TTTTTCAAA TATTTTCATTG	180
	CGATCCAAAT GTAATAATGC CTCCTTAATG CATAGTTTCT ATTTCTTCTA ATTTATATTC	240
50	AAGTGGCTGT TCATGTCCCT CAAGCTTCAC CTGCATAGAA ATGTCTAATA TATTTAAAGC	300
	AACTACTTTC CCGGTTACCA TCAGGCGTTT CAATTGCTTC ACCAATATCA GGTAATTGTG	360
	CACGTACTTC CTCATAATAG TCATTTTCAT ATTTTAAACA ACACATCAAG ACGACCACAT	420
	GCAACCAGAA ATTTTAGTTG GATTTAATGA TAAATTTGGG TCCCTTAGCC ATCTTTAATC	480
55	GATACTGGGT TCAAAATCCC CCTAGAAAAT GTAGGAACAA CATAACGACC TACCCACAAG	540

# EP 0 841 394 A2

GGTCCCGATA CCGCCAAGCA ATTGGGCTTC AAT

573

## (2) INFORMATION FOR SEQ ID NO:49:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 870 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

CCTTAAGAAG	CACCTTTACC	AATTCGCGCA	AAATCTGTAA	ANTCAAATGT	AGTAGGTAAT	60
GTTTTTTTAG	GTTGAACCAT	TTCTTCTAAT	TTAAGTAGTC	TTGCATCTGN	CACTTCNACA	120
ATACCCACAT	TTGGGTCNAT	AGTTGCGANC	GGATAGTTTN	CTCCCAAGGC	ACCCGCCTTT	180
GTAATTGCAT	TAAATAATGT	TGATTTACCA	ACGTTTGGTA	AGCCAACGAT	ACCTGCTGTT	240
AAAGCCATGA	ATCATTCCTC	TATCTTNGTG	TATCATCCTG	AGATACGATG	ATTTTTTTTA	300
GTTTTTTATC	AAACGTTTGA	CGTGGAATCA	TAATGCTTCG	TTGACAATTT	TCACACTTAA	360
TTCTAATGTC	AGCACCCATT	CTAATAATTT	TAAAACGATT	GGTCCACAC	GCATGTTGTT	420
TTTTCATTTT	TACTATATCA	TTTATTCCAT	ATTGNGACGC	CATTAAATAT	GACCTCCATG	480
TATTATAAAC	TACTAAACAC	CGNTTTCAC	ATGATTAAAA	GGTGTCATTA	TTGGNTGTGG	540
TGTCTTGATA	CCTTCTTGTA	AAAACATTTT	GTGGACTTCT	TTACGAATAA	TACGCGCACC	600
AGAGAAACCN	TCACCTGGTA	TTGTTTCTGC	TGAAACCCNT	AATATTACTC	TAGTATCTTC	660
AAAGGCATCA	ATACCAATCA	CAACAGGATC	ACTTACAAAT	AAATNNATAT	TTACTACGCA	720
GAGGATGGTA	AATAATGTGT	TTAATTTCCT	CTCAACATTA	TCAATATTTT	CATCTACAGA	780
TACTGGAATT	TTAACAATTG	CTGGTTCCAC	TCGGNAATTG	AGTAANTCGG	TATTCACCCC	840
NTNCTACTAT	TTGGGAAGAT	TGGRDCTNS				870

## (2) INFORMATION FOR SEQ ID NO:50:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 519 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

TAAGATATTA	AGTTATAAAA	ATGATTTGAA	AAATAAAGTA	GAAAACTTAA	ACAATTTAAG	60
TCCAATAAAT	ACNATGTTGC	GTGGATATGC	AATTGTTAAT	AAAAAAGACG	AAGTCATTAC	120

EP 0 841 394 A2

5  
10

GAGTACTAAA	GATTTAACAG	AAAATGATCA	ATTGACGTTA	ACAATGAAAG	ATGGCTTAGT	180
AGATGCAAAA	GTTACGAAAG	TAAGGTGTAA	TAATGACTAA	AGAAACGCAA	AGTTTTGAAG	240
AAATGATGCA	AGAATTAGAG	CAAATTGTTC	AAAAATTAGA	TAATGAAACA	GGTATCTTTA	300
GAGGAATCAT	TAGATTTATA	TCAACGTGGG	TNNTGAAACT	ATCAGCAGCT	TGGTGACACA	360
ACTTTTAAAA	AATGCCGANA	AAAAGGNGAA	TGACTTAATA	AAAGAAGAAG	CTGAGATTGT	420
TAAAAATGAC	GATCTTCCGA	NTGANTTAAT	TATAGANGGA	NGTCAATATG	ATNTCCGGTG	480
GCGATAATTA	NCCCAGNAAT	GGTTCTCNGT	NGGAAGGAG			519

(2) INFORMATION FOR SEQ ID NO:51:

15 (i) SEQUENCE CHARACTERISTICS:

- 20
- (A) LENGTH: 734 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

30  
35  
40

CCAATCGCNT	GTNTGACTGT	NTTGTCAC	TG ANAGCGTTAT	CTTNTTTACC	AGANTTGACA	60
ACTNACATTN	TCGTANTCAT	AGGTTGACTA	CGCNTAACTT	GATAGTCACC	TGTATCTTTC	120
AATTGTTTTTA	AAGAGTCTTT	GNCTAAGCTA	TCCGTACCTC	TATCATCTGT	GAGCCANAGT	180
CGTTTCACCT	TTTTTCATTG	ATAGGAATGC	TGTTTCACCA	GCAGGCATTA	CTTTTGCTTG	240
TACTTTGTTA	AGTTAAGACT	TTTCGCCCCA	GTATTGATCA	TTTTTGTTAA	AGTCTGCAGA	300
CTCATCTTTT	TTGTGTTTAC	CTAATTTAAA	TGGACCAGTA	CCATCGAACT	TTTTAACGCC	360
ATCTTTTGTT	GTACCGTTTT	TAAAGTCTTT	TGGAGACACA	AATACATATG	GACGAGGCAT	420
CGCTAATTCA	GCCAATGCAG	GTTGATATGC	TTCTTTCAAA	TTCAATTCAA	CCGTGTACTT	480
ATCTTTAACT	TTAACATTGT	CAATTAATGT	TGAAATCTTT	AACCAAGAAT	GCAATTTTTT	540
ATTTTGTTGA	NCCGCATCAA	TATTTTCTT	AACTGCGTCA	GCATCAAATG	TTGNTCCATC	600
ATGGAATCTC	ACATCATCTC	TTAAATGGAA	TTGTATTAAG	TCTTCCCATC	TTCAAGACAC	660
ATCCCATTTT	TTAGCTAGGT	AAAGGCTTAA	TACCATCTTT	CGGTGTTACG	GTACAAAGCG	720
GCTCGTARRD	CTNS					734

45 (2) INFORMATION FOR SEQ ID NO:52:

50 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1219 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

55 (ii) MOLECULE TYPE: Genomic DNA



EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

5	GNGNTNCTTA	TNCGATTAC	AGTGAAGTAG	GTGATAAGAT	GATTGGTCCC	AAGCTGAAAT	60
	CCCCAAAATT	GTGCCAATTG	ACTATATTTT	NCAAACAGGC	GATATTGTTG	AAATACGTAC	120
	TAGTAAACAT	TCATATGGAC	CAAGTCGTGA	TTGGTTGAAN	ATTGTTAAAT	CGTCTAGTGC	180
	CAAAGGTAAA	ATTAAAAGTT	TCTTCNCAAN	ACAAGATCGT	TCATCTAATA	TTGAAAAAGC	240
	CCGAATGATG	GTTGAAGTTG	AAATAAAAGA	CCAAGGATTT	AGAGTCGAAG	ATATTTTGAC	300
10	AGAGAAAAAT	ATTCAGGTTG	TTAATGANAA	ATATAACTTT	NCAAATGAAG	ATGATTTATT	360
	CGCAGCTGTA	GGATTTGGCG	GCGTGACATC	CTTACAGATT	GTTAATAAAT	TAAGTGAAG	420
	ACAACGTATT	TTAGATAAAC	AACGTGCTTT	AAATGAAGCA	CAAGAAGTTA	CGAAATCATT	480
	GCCTATTAAA	GACAACATCA	TTACTGATAG	TGGTGTCTAT	GTAGAAGGTT	TAGAAAATGT	540
15	ACTTATCAAG	TTGNCAAAAT	GTTGTAATCC	TATACCAGGT	GATGATATTG	TAGGTTATAT	600
	CACCAAAGGT	CACGGCACGA	AAGTACATCG	CACTGATTGC	CCAAATATTA	AGAACGAAAC	660
	TGAACGACTA	ATTAATGTTG	AATGGGTAAA	ATCNAAGAGC	GCAACTCAAA	AATATCAGGT	720
	TGATTTAGAG	GTAAGTGGT	ATGACCGAAA	TGGCTTGTG	AATGAAGTAC	TACAAGCTGT	780
20	TAGCTCGACA	GNCGGCAATT	TAATTAAAGT	TTCAGGACGT	TCAGATATTG	ATNNNAATNC	840
	AATAATAAAT	ATTAGTGTCA	TGGTGAAAAA	CGTGAATGAT	GTTTATCGTG	TGGTAGAAAA	900
	GACCAAAACA	CTTGGTGATG	TTTATACAGG	TAACAAGAGT	TTGGGAACCTA	GAGGGTCCAA	960
25	AAATATTGAA	AGTAGGTGGT	ACAAAAGAGG	TTAAAGAAGC	ATCGGGTGGG	CGAATTGATA	1020
	CATTAATAAA	TCAAATCAAA	AAAAGGATAT	TGGTTTATTA	GTGCGGGGATC	CGGNCAAGAA	1080
	CTCTACAAGA	GCAAAGATGC	AGATGTAATT	GCAAAGAAAA	TTGCTAATGC	AAGATTATTT	1140
	GAAGATGACA	ATAATAAATT	AACTTTAAT	ATCCAACAAA	TGAATGGTGA	AATACTATCA	1200
30	GTTTCACAAT	TTACTCTCG					1219

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 498 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

	ACTAAAAATGA	TTAAATTACT	TATCTCTTTT	CATCTGATTG	ATTATCGAAA	TTTCTTCCTT	60
	CTAAACCTGC	TAAGTCTTCT	TTAGAAGCTG	NAGGTGCTTT	CATTTCAAAT	ATCTCATTCA	120
50	CTACTGTGTA	ATCGGAATAT	CCTAATCTGG	CAATAGGTTT	AATCGACTTA	ATGTCCAATT	180
	TACCATTATC	AAGAATAACC	NTATCGGCAA	TATGAACTTG	GGCAACTCTT	CCTATAACAA	240
	TATCTACGGT	AGATACTGGG	TCTCTGGCG	GAATACGAAT	CGGGTGAACG	GTAATCACAT	300
	NCAAATGTGA	CTGGCGATCT	TNGACTCGAT	AACCNCGGGG	CTTCTATACA	ATGNTNCTTN	360
55	CGGGCACAA	CGGNATATTN	AAAATNCAAC	CCCCNCAGG	TGGGTAGGTG	CCTTCCGACG	420
	GATAAATTAA	CAGGCTTCTC	TTTAAATCAT	ACGGTGGCCA	TGTNCCAATA	CAAACCCACC	480

CCGGTTTCTT CCTGGAAT

498

## (2) INFORMATION FOR SEQ ID NO:54:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 922 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

CTGTCAATGA	ATATGATGAT	ACTGATAAAA	TTTCTCTAA	CCCATCAAAC	AAGAAAACAG	60
AAGATTATAT	TTCAGGAAGG	TTTGGTTGAT	ATATAATGGC	AATAATTAGA	CAGCGATATC	120
AGGAGCAACT	TGATGATTTA	ATAAAGGAAT	TACGTCGGTT	AGGTGCAAAT	GTCTATGTGA	180
GTATTGAAAA	TGGTATAAAA	TCATTAAGTA	TTGACGATAG	AGGCTTTGNA	CGACAAACAG	240
TTAAAAACGA	TAAACATATC	AATCAATTAA	ATTATGATAT	TAATGAGCGA	GTTATCATGT	300
TAATTACAAA	GCAACAGNCC	ATTGCGAGTG	ATTTGCGTAT	GATGATTTCT	TCAANTAAAA	360
ATCGNCTCCG	ATTTAGAAAAG	AATAGGAGAT	AATGCATCGA	GTATTGCCAA	TATTCGATTG	420
CGTACAAAGA	TTACAGATGA	TTATGTGTTA	ACCCGTTTAA	AGACAATGGG	TAAATTAGCT	480
ATGTTAATGC	TAAAGGACTT	AGATCAAGCA	TTTAAAAAGA	AAGATACCGT	ATTAATAAGA	540
GAAATAATTG	AGCGTGATGA	AGATATCGAT	GACTTATATA	GTCAATTAT	TAACGCAACG	600
TATCTTATTG	ATAACGATCC	ATTTGTCGCT	GCACAAGCTC	ATTTAGCAGC	AAGACATTTA	660
GAACGTATTG	GTGATCATAT	TATTAACATC	GCTGAAAGTG	TTTATTTTTA	TTTAACAGGT	720
ACACATTACG	AACAAATAAC	TTAAAGTTAT	TACTATAAAA	TCCCTTACGA	TAAATATATA	780
TTTCTATCAT	TCATAAACCC	TCAAAAAAAA	CCAAGATTCT	CACAAATTAG	NAATGNGTGA	840
AAANCTINGN	GNTATATTTT	GGTTCTTACT	TATTAAATNG	GTCTCGCATC	TTAGGNTATT	900
TGGNTTGGTC	AATTCATCT	TG				922

## (2) INFORMATION FOR SEQ ID NO:55:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1209 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

CCAATCGGTC	TTATCTTTCA	ACACGTTTGA	TTGTANCGGA	TATCACCCGG	TTCATAAATC	60
------------	------------	------------	------------	------------	------------	----

EP 0 841 394 A2

5 CNAATCGTT GTAATATATT TCTCTTTATA TTCATTATTT CTAGGAATCC ACCCCCNCCG 120  
 TCGGGCAAG TTTCTGGGAA ATTTAACAGC TATCCAGCCT TTCAACAAAT AAATCTGAAA 180  
 TTCGATGTTT TAAAATTTCT GCTTCTTGAT GACTTCTTC CCAATTATAT TTCAATATTT 240  
 CTATTAAAAA TAATCTAAT AGTCGATGTC TCTTAATGAT ATCAAGCGTA TCGGTTAAAC 300  
 CATCCTCTGT TAATCTAACA CCTTGTATG GTTTTGTTC AACATAGCCT GCTTTTCAA 360  
 GACGTCCTAC CATTTCACTT ACAGATGGAG GCTTAATATT TAAAAATTGA GATAAGATTT 420  
 10 TATTTGTCAC AAAGTTTFTA TCGCCATTAA TTCGTAAGGA TTTGCCTTTA AAATAAGTCC 480  
 TCTTTTCTT CAGNTAACAT ACTTTCACCT CAACAGACAT TGTCTATATT ATATCAGGAA 540  
 TTTACTTGAC ATGATAAATA TTCTCAGTTT ATTANACAAT TAATTAGGTT AGCCTAAACT 600  
 TTTAATTAGG AGGTATAAAC GTTGTGTTAGA AACANAAGAT TTAAATCTGT TTTTAGGTAA 660  
 15 TAAGCATGTA CTTAAAAACA TTCTCTATC GATACCAGCA CGAGCGGCGA AATAATTGGT 720  
 ATCATGGGCC CGAATGGTGC TGGTAAATCT TCCCTTATCA AGTCTTTAAT TGGTGAATTT 780  
 AATGCTACCG GTACTAAATN GTTNTATAAC AAACCTATAC AACAACAACN GCAACAGATT 840  
 ACATATATTC CACAAAAAGC ACATATTGAT TTAGATTTTC CTATAAGTGT GGAACAAGTG 900  
 20 ATTTTATCAG GTTGCTACAA AGAAATTGGA TGGTTTAGAC GACCTAATAA ATCAGCAAGG 960  
 GATTAAACTC AAACAGTTAT TAAGCGATTT AGAATTAGAA TCTTTTACGT CATCGACAAA 1020  
 ATTTCAAGAA TTAAGTGGGT GGGACAAATT ACAACGGTGG TGCTTANGTA GGAAAGAGCA 1080  
 TTGGATGGTC CCGAAAAGTG AAGGTTTATT TTCTTTAGAT NGAGCCCGGT TGGTCGGGAA 1140  
 25 TTTGGATTTT AGGGTAGCGA AAAAATTAAT CAATGACAAA AAATCCGAGA ACTTTAAAAC 1200  
 AACAAAGGA 1209

(2) INFORMATION FOR SEQ ID NO:56:

30

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 369 base pairs
- (B) TYPE: nucleic acid
- 35 (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

45 GAATTCATNT TATCTATCGG NAATCATTTG TAATACATCT ACTAATTATA GTGTTTATT 60  
 CGATAATATA ACTGCATGTA CATCTANAGA AGCCTCTTGC CTNGAAATCG CACCGTCAAA 120  
 CTTAAACTA TAACCTATTA AATCTGCTTG NGTGNCATTA CCACTCACAG NACAGCGATT 180  
 AAATGATGCA GTAGCAACCA AATCGTTTCA TGCACTTNCA CATAANTACG ACTGACATTA 240  
 NGCTGTGTAG ATGTACCTGC TTCTANCTCG CCAAGAACAT CTTGGANTNC GNGNTNGCTA 300  
 50 TTTGGNGCAA TGTCANCTTC ATCCATNGAG CGCTCAATAG TGTCAGCTGC CAGAGCGGCA 360  
 TAACTGCTT 369

(2) INFORMATION FOR SEQ ID NO:57:

55

(i) SEQUENCE CHARACTERISTICS:

# EP 0 841 394 A2

(A) LENGTH: 268 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

CCTTATTATA AATTAAAGAA AGGTAGTGGT TTCTATGAAA TACAATACTA ATGTTAAACA	60
TACAACCTTA GAAGCGTTTG TCACAACGTG CAATGATTG GGTATTGAAT TAATTATCAA	120
TGAAGCACTT CGAGAGGTAA GAAAACGACA GCTCATAGAA CTTATAGATG ACGCACTCGT	180
CAATAAGAT GAAGCAGCAT TTAATCAATA TACGGCAGAA TACAAAAATT TGGAGGCATT	240
TCTCGGTGNA TAACATTGAT TCGAATTC	268

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 512 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

CTATACAACCT AAAGTAATTC ACAATTCAAA CTTTGTGAAAG GGTGTACAAA GTGAAAGTAG	60
CAGTCATTAT GGGCAGTTCT TCCGATTGGA AAATTATGCA AGAGAGTTGT AACATGTTGG	120
ATTATTTGGA AATTCCTGAC GAAAAACAAG TAGTATCCGC ACATCGTACG CCAAAAATGA	180
TGGTTCAATT TGCTTCTGAA GCGAGAGAAA GAGGTATAAA CATTATCATT GCAGGCGCTG	240
GCGGTGCGGN ACATTTACCA GGTATGGTTG CATCATTAAC GACGCTACCA GTTATTGGTG	300
TGCCGATTGA AACANAAAAG TTAAAGGGT ATTAGATTCT TTTATTATCA ATTGCTCAAA	360
ATGNCAGGAG GGTATTCNG NTGGAACGAC TTGCAANTGG GTGCAGNAGG GTGCTTAAAA	420
ACGCAGGTAT TCTTGGCAGG AAGAATGTTA AGTNTTCAAA ANTCTCTTT TGGNNGGANA	480
ANTTAATCAA GNTTGATCTT CCGNAATCAA AA	512

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 321 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

# EP 0 841 394 A2

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

GAATTC	CCCA	AGTTC	CTCGC	ATTTA	ATTGA	ACTTT	TATAAT	ATTTG	TAGTC	ATCCA	CAAGC	60
AAATAT	TATAC	ACTCC	CAAAG	GTTTA	CATAA	TTCCA	AAATTA	GACAC	ACAAC	GGTGG	ATTGA	120
AAAAT	GGCAA	AACCA	TGGC	AACA	TATCA	ATTGG	TTAC	TTTGT	ATTGG	TGAAAA	AAAT	180
AGATT	GTAGT	GNTAT	TGGTA	TTTGT	GGATA	TGAAT	ATCGA	CAATT	AAAAGC	AAGAA	CAGN	240
ACTAA	NTTTA	TTTNAT	AAAT	TACAT	CCAAG	TTTGT	GNGGA	CAAGG	GGTAC	GCATG	TGAGG	300
CTNTT	ACAGC	AATCA	CAAAT	T								321

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 422 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

GGCACG	ACTC	GTGCC	GCTCC	TGCAC	AAGAT	GTAGG	AGCAT	TTTTA	TATTT	AAATA	ACTAG	60
AGTAAT	TAAAC	GTAAG	GCGT	GTGAT	ACAGT	GAAAA	CAATT	GATTAA	AATTA	ACACCG	AAGC	120
AAGAAA	AGTT	TGTGC	TAGGA	CTCAT	AGAGG	GCAAG	AGCCA	ACGGA	AAGCA	TATATT	GACG	180
CAGGGT	ATTTC	GACTAA	AGGT	AAAAG	TGATA	ATTAT	ATTGA	TAGCC	GAGCT	TTGAG	TTGA	240
GTAAGA	ATAG	TGCGG	GTTTA	GATAG	GTATG	AAGAAT	TGCG	TCAAG	AAGCA	GGTTG	AACAA	300
TCAAA	ATGGA	CACGC	CAAAA	GGGTT	TTGAA	GAATAT	GAGT	GGGTAA	AAGAA	TGTNG	CTAAG	360
AATTG	ACATT	GNAAT	AGAGG	GGAGG	TAAAG	GAAAG	GGACA	GGTTG	ATGCA	TTCCCC	CCTA	420
GT												422

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 395 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

# EP 0 841 394 A2

	GGCCCGAGCG	GGACTTAAAA	AGGCCACCAC	TGGTTGTGAC	CTANCCCTTAT	TTACNTTTAT	60
5	AAATATAAGG	AGGAGGTAGT	AGTGAAAGAC	TTATTGCAAG	CACAGCAAAA	GCTTATACCG	120
	GATCTCATAG	ATAAAATGTA	TAAACGTTTC	TCTATTCTTA	CTACTATCTC	AAAAAATCAG	180
	CCTGTCGGAC	GTCGAAGTTT	AAGCGAACAT	ATGGATATGA	CTGAACGTGT	ACTGCGCTCT	240
	GAAACAGATA	TGCTTAAGAA	ACAAGATGTG	ATAAAAGTTA	AGCCTACCGG	AATGGAAATT	300
10	ACAGCTGAAG	GTGAGCAACT	GATTTTCGCA	TTGAAAGGTT	ACTGNGATAT	CTATGGAGAT	360
	GATAATCGTC	TTGTCAAGAA	GGGTATTTAA	AGAAT			395

## (2) INFORMATION FOR SEQ ID NO:62:

15

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1493 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

### (ii) MOLECULE TYPE: Genomic DNA

25

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

	CTGTGATTTA	TCATTCCGATT	GCATGATTAG	AGGGAGGGAT	TAAACGTGAC	ATATCATGAG	60
30	CGTGTTTT	CATTAAGAGC	AGANAGTAAN	AGNNCCGCAT	TTGATTTTCG	ATTCGGNAGA	120
	TTTATTTACC	NAAGAAGAAT	GGCTAAGTAT	GTCTCTTGCA	GAAAGACNCA	AAGCTGANAA	180
	AGCATTTCGA	CACGNAGTTA	AAAATATGGA	CGATGTAAGA	ATGCCCTTCT	CAAGTGTCCA	240
	TGACGCCCCA	AGTAAAATTA	TATAATGTTG	TATATTCTGA	TANCGGCATT	AAACGTAATT	300
35	TTAAACNAGC	TGAAAATGAA	GGATTCTAAT	ATCATTTCGT	TTATATATAG	CAGACATGAT	360
	AGAATTTTAT	ATGTAAATCT	TGTAGGTAAT	CGTTTTNNAA	ATAATATAAG	TATGACTAAA	420
	NCCACATCCA	ATATAGGACG	GGGCTTTTAG	TATCGTTCCA	GATTGTGGG	GAGTTAGGGT	480
	TTACTTAAAT	AAAGGGGTCT	GCCCTCAATT	GCTCACCAAC	GAGGGCAAGT	TACATCAATC	540
40	ACACTTCAAT	TGCCGCCGAC	TAGGGGTAGT	AATCATTGGC	AATAAGAACT	AGTTAGTGAC	600
	TAACGGATTT	ACGTTCCATA	AGCAAAGTGA	TACAAACGCT	CAGCATCAAT	GTAAGGCATA	660
	GAATCAATAG	TTAGGTAACC	ATCTTTGATG	TCACCAATTA	ATCCTTGATT	TACACCAAGT	720
	TTACAGTAAA	TATCATAGCT	ACTTACGTCT	GAGAAAATAA	GCGCTTCAAT	CTCTTTTCTC	780
45	ATAATCATCA	CTCCAATGTT	TATATATTAT	TTATATAAAC	TCTCTCGTCT	CTCTCTATTT	840
	ATTAACCTCT	TACAAGTCTA	ATATTACATG	AATTTCCAAA	TAAATAAAGA	GGTTTGTGAT	900
	GTATTTTACA	AATTATACAC	TATTTTGGAA	AATTAAGAAA	TAGTTAATTA	TATAAAAGTT	960
	TAATAAGTCA	GAAAATATGA	TAAAATGTAG	ATGTTCTTTA	GACATTAAAA	GCTTCTAACA	1020
50	TGATACGGAA	TATGAGGTTT	CTGTATCACA	TTAGAAGCTT	TTTTGTTGCG	GTGCTTATAT	1080
	TAATAATGGC	ATGAATGAAC	GCATGCTGTC	TTAGTTTCAG	GAATCGGCTA	GAATATTAAA	1140
	CTTCCGTTCT	TTGAATATTT	NTTAATTTCT	TACTAATATC	ATCCAGTTCA	TTCTTCGCCCT	1200
	TTTTAACCCA	ATCTCCTTGA	GCTACATTAT	CAAATTCACC	TTTAAATTTT	TTGANATGGC	1260
55	TACCTATATT	TCACAGCTCA	CTTTAGATAT	CGAACTTGTT	TCTTCAGAGT	TACTTGTATAT	1320
	ATTACCAGGT	TTTATTTGAT	ATTCAAGTTG	ACTCATAACT	TGTATCCTCA	CTTGAAGTAT	1380

# EP 0 841 394 A2

AGTGTACAGA ATCTGTAAAA TAAATGTTTT CTCTTTTATT TTCTACAAAG GTAAATTCTA 1440  
GATTTTTTGA CCCTACAGAT GAGCCTTTTA AATCACCATC NCCTTTTATT AAG 1493

## (2) INFORMATION FOR SEQ ID NO:63:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1215 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

TGATGTGATC	CAACCAAGTG	ATAGCAATGT	ATATGCAGTT	GGAGAATGTG	CAGAACATAC	60
CGCCAAAGTT	TATGGACTCG	TTGCACCACT	ATATGAACAA	GGTAAAGTAA	AGCTGATCAT	120
TTAACAAATA	AAGAAACGAA	CGGATACAAG	GGATCAACAA	CAAATTACGT	CATTAAAAGT	180
TTCTGGTTGT	GACTTGTATA	GTGCTGGTCA	AAGTTGTAGA	AAATGCAGAA	ATTAAAGGTA	240
TTGAAATATT	TAATAGTGTT	GATAATAACA	ATAAAAAAAT	CTTTTAAAAA	GACGGTAATG	300
TAGTTGGTGC	AGTATCGTGC	CAATATGGTG	ATATCGATGA	TGGTTCACGC	TTTTATAACA	360
TGATGAAAAA	AGGTGAATCC	ACTGAAGATT	ACACACTTGT	ATCATTCGCT	ACTAAAGGTG	420
GAGAAGAGGC	ATCGCTATCA	ATTGCTGATA	TGGCTGATGA	TGAAACAATT	TGTGGTTGTA	480
ATGGTGTTGA	TAAAGGTACT	ATAGTAAATG	CGATTACGGA	AAATGGCTTT	ACAACAGTTG	540
AAGAAGTAAC	GGCTAAAACA	AAAGCGGGGA	ATTGATGTGG	TAAATGTAAA	CCGCAAATTG	600
CTCAATATT	GCAGCACACC	TTAGGAGATG	ACTTTGTTGC	CGCAAAACCT	GCTGGTATAT	660
GTGGTTGTAC	TGATTTGACA	CGCGATCAAA	TTGTAACGCA	AATAAGAGCG	AAAGGTTTAA	720
AAACATCTAA	AGAAGTTCGA	CATGTTTTAA	ACTTTAAAAA	TAAAGGTGGA	TGTCCAAAAT	780
GTCGACCAGC	AATCAACTAT	TATTTAAACA	TGGTTTATCC	ACATGATCAT	GAAGATGAAA	840
GAGAATCAAG	ATTTGCTAAC	GAACGTTACC	ATGCGAATAT	TCAAAATGAT	GGTACATTTT	900
CTGTTATACC	TCAAAATGCGT	GGAGGTGTTA	CAGATGCAGA	CCAACTGATT	CGTCTAGGAG	960
AAGTGGCTAA	GAAATATCAT	GTGCCACTAG	TTAAAGTGAC	AGGTTCGCAA	CGTGTGTTGGT	1020
TGTATGGAGG	TTAAAAAAGA	AGAATTACCA	AATATATGGG	AAGACTAGGT	ATGCGTTCAG	1080
CATCAGCTTA	TGGTAAGAAA	ACACGCTCAA	GTTAAAAGCT	GTGTTGGGTA	AAGAGTTTGG	1140
GCGATTGGGT	ACGCAATACA	ACGACAACGA	CTTGGGATCC	GTTTAGAAAA	ACATTTGGGT	1200
TCAATCGATC	AACCC					1215

## (2) INFORMATION FOR SEQ ID NO:64:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 760 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

5	TTACCTCCAA AATCCTTTTA AATCCATGCC CATAGAAGAT GANNCCCAA ATACATTANC	60
10	TAAATCGAAA CATTGCCATC CATAACTACC GTCGAAATCC CATCCTCTAT TCTCTAATGT	120
	TCTCAAATAA NTTAATGTAC TGCTATTACT ATTACTTTTA TTATTTGAAG ACACTGTTTN	180
	TGGTNTTGGT TCTACTAATG GTGTCAATGG CACTTTTAAN TNTTGACCAA NAAATANTAA	240
	ATTAGGATTT GCTATATTAT TTGTATTTTG AATATTGAA ACTGAAGTTT NGTACTTTTA	300
15	ATGCNATAGC ACTAAGTGTG TCTCCTTTTT TTACAGTATA GAGTTGTGTT TTTGGAGCTT	360
	CTCTAAACCT GTAGTAACCA AAGTAATTAG TAAGTAACNT TAATCGTTTT CTTCTGATAA	420
	GAATNTTTGA GCTTCCAAGT TTGCAATNTT AANTTCTTNA GTAGGTNCAT TGTANTTAA	480
	TAACTAGATT GTTACCNTGG CTTGAGTTTT TCGGAGTATT TGAAATCTTT ATATCTTGAT	540
20	TAATTTCAAT TCCGTTTGAA ATTGCTGATT TGTTGTCTAA CTTTAACTT GTGTCCGATG	600
	TTTTAACAGC ACCTTCATTT TTTATTTTGT CTTTGTCTGT ATTTTATTA GCATTTAACT	660
	CTGATTTCGC TCGAATACAT TTTGCTCATA CCCTCTGTGA AAATCTTTAG ATTTATCAAT	720
	TTCATCTGCA TATGCTTNGT TCGACATACC CAATGCCAAG	760

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 846 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

40	ATATGCTGAA AATGTATCAT CAGGATTCAA GATATCAATT AAATGATGTG GAATACCATC	60
	CATTTCTTCA GGTGTACTT GTGCAGTTCC AATATTCATA TGTTTGTAAG CTTGCATAGA	120
	ATCACCGCTT ATGATTTTAC CATTGATACG CTTGCGAGT TCTATGCTAA GCTCTGTTTT	180
45	ACCTGAAGCA GTTGGCCCCA CAATTACTAC AATAAAAGGC TTATTTTAT TCATATTTAA	240
	TTCCACTTTC TTTATTTTT ATAAAATTGA TTCAACCATT TAAACAAGTC TTCCCAAATT	300
	GTGTCATGAT CTTTTTTTCA ATAACACTTC ATGACGTTTA TGATGATAAA GATTCACTTT	360
	TAACAATTTT TAATGCCAGC TTCTCTATAT ATATTCGCCA ATTTTAAAT CCCTTTACCA	420
50	TAATCACCTA AAGAATCGGC ATATCCTGAA ACTAATAATA TAGGCATGGC ATGATTTAAC	480
	ATTTTCATAT TTTTAAATTG TGATGTATGT AGCATATAAT AGGCTGTGTG ATATAATAAT	540
	TGATNTGACA CATTTAAATCC ACTATATGGA TCTTTAATAA AGTTATCTAC TTCAATTGGA	600
	TTACTAGAAA TCCAATCACT NTGTGTACGT AATGGACGTA TTTTATTATT AAAACTATNT	660
55	GATACTAACT GGTAAACCCA TNCAACTCGT TTCTCAGCAC CATAAATTTG TGTAATCAGT	720
	TGTAACACTT TAACGGTTGG GTAAACCTTT CCATAATTGA ATACATACCA AGGTCACCAC	780



EP 0 841 394 A2

TTTAAATTA GACCATCAAC ATATTGCGG ATATTGGTTT CTTACAAAAC AATCTTAGTT 840  
ATACGG 846

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2854 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

AAGNTGAATT	TTAGTACAA	TTTCCACTGT	ATTTATCTGT	TANCCACTTT	ATACCCGCAA	60
TATTTTCGTTG	ATCTTATATA	AACACTTTTG	GCTGCTTTTC	CTTGACCNAG	GNTAANATTA	120
NAGATTTTCC	TTGTAAAGAG	TGTGTTTTGT	AGTATCATAA	ATATGTAANC	TGTATNAAAN	180
TCCACCACAT	TTAGGACCT	GGTGANAGAT	GTCGNAAAAG	AANATTTTAA	TATTANCTAG	240
TATTATGTTA	ATCATANTAA	TTAGTGTTGC	AAGTATCNAN	TTTAAATGA	AATATGACGA	300
AAAAGAAAAC	CAAAAATCAA	TTTATTATAA	AGAACAACAA	NCGCGCATT	CACTTTATCT	360
TAAGCATAAT	ACTATAGAAC	CGAACACAAC	CAANTCTGNA	CATTTNCNA	AATTGGAANC	420
AANTCCTATG	GGAAGTGCTG	TGATTGAAGG	TTACATAAAT	GANAAATAAG	AAGATGATTT	480
TGTTGCCTAT	GCATCACCTG	AAAATAACTT	TCAATTTGTA	GGTGATTIAN	TTNNAAAGTG	540
AAAGATTAAG	TGAGTTACTA	AAACCAGCGC	ATCAATTNNA	ACATCNCCAG	ATGATATNNN	600
NNAAGAACTA	AATANNAAGT	AGAGTCACTA	ANGTTAGGAG	TTACTTTAAT	GAACAAAAAA	660
CATGTTTTTG	TAATTATTGG	TGTCATTTTA	TGTATATGTA	TAGTTGCATC	AGTCATTTAT	720
TTAAAAGTGA	AATATGACGA	AAAAGAAAAA	CCAAAAAGCA	AAATACTATA	AAGAACAACA	780
GGAGCGTATT	ACACTCTATC	TCCAGCATAA	TACCAAAGAA	CCCAATACCA	TCAAATCTGT	840
GCATTTCAACA	AGTTTAAAAA	GAGGACCCAT	GGGCGATGCC	GTAATTGAAG	GCTACATCAA	900
TGAAANCAAA	AAAGATAATT	TTACTGCTTA	TGCTACACCA	GAACATAATT	ATCAGTTTGG	960
TGGTGCTATG	ATAGAAAAGT	AAAGATTAAAG	TGAGTTACTA	AAACCAGCGC	ATCAATTAAA	1020
ATCACCTGAC	GAAATCAAAG	AAGAATTAGA	CACCNAAGAA	GGCCACTAGG	GTCTTCTTTA	1080
TTTTTGATTT	AATCTTCCAA	TAATCTATGT	CATTGCTATC	GAAGGTGTTT	CGCAATTAAT	1140
ATCAATCACT	TCATCATCAC	CAATACTTCC	CCAGTTTGT	ACAGTACATT	AACACAAGCG	1200
AACCCCATTA	ATGTAAATGA	AATAATAGTT	TAGCCATTAT	AAAAACATTA	TATACCGAAT	1260
AACAATGGCG	AGGACGTGAT	TTATTGACNA	CTATCAAATT	TTCTCANATG	TGTGTGATTT	1320
TTTATAANNN	AGGGTAGAGC	AAGGTATAAT	ATCCTTTCAA	TCGGTTTTCA	TATTTTATAT	1380
ATTTTTTAAA	TATAAGCGCT	AAATGTTTTA	ACTAAAGCAT	AGATTGACAA	GATGTTATAC	1440
AGAATTTTAA	ATTCTATCCA	ATATTGTTTCG	AAGTGTAAGT	TCAGTGGANT	GGTATTAAAC	1500
AATGTAAAGG	AGAGATTGCA	NATGCCGTAT	AATTACAAGA	AACACAATGG	AGAGTTAATG	1560
TCTGTAATGA	GCCAAGGTGA	AAAGTTTATT	CATCCAATCC	ACCGGTGAAT	GATGAACCTA	1620
GTGCATTGAT	TAAAGCTATT	AATCTCCAAC	AAATTAACGG	TTGTCATTAT	TGTGTTGATA	1680
NCCANAANAA	AGAAGTAAAG	GAANTGGGTG	TAACACAAC	GANNATTGAT	GAAGTCTTGA	1740

# EP 0 841 394 A2

	NNTTTAGACA	TTTAGANTTA	TTTACTTGAT	CAAGAAAAAG	TGACGCTTGA	AGTTGCAGAA	1800
	ATGNTAAATT	CAANCAAAGA	CNTTAAGAAG	TTTGAAATTA	TTGACCGGCN	AAAATCATTT	1860
5	TATGATGAAG	AACAANTTAT	TGATCTTGTC	TTTGTGTAA	ACCAAATTAA	CGGTGGAAC	1920
	AGACTAAATA	TTATTAGTGA	TAGACTATAA	TTGTTTCATAT	AAATGCAGAG	TTTCATCTCG	1980
	AACGCTATAT	CATAACAANT	CATGCCACTA	TACAGGTCAA	ATCTTGATA	GTGGCATT	2040
	TAATTTATCC	CTTTGAATAC	TGTTATTTAA	CGAATATCGG	TCCACCTGGT	CCAACAGGGA	2100
10	TACCTAATAA	GAACCAAATG	ATGACAAATA	CTGTCCATAC	AATGCTTAGC	GCGATTGAAT	2160
	ACGGCATTAA	ACTAGAAAGT	AAGGCTCCGA	GTTTCATGCG	TTTATCGTAT	TTTTGTGCAT	2220
	AAGTTAATAA	TAAAGGTAAG	TACGGCATCA	TCGGTGTAAT	TGGATTGGTA	ATTGAATCGC	2280
	CTACACGGTA	AATGACTTGC	GTAAATGCGG	GATGAAAGCC	GATAAGGATT	AACATTGGTA	2340
15	CGAATATCGG	TCCTAAAATA	CCCCATTTAG	CCGATGCGCT	TCCGATTAAC	ATGTTGACCA	2400
	TTGCACTCAG	TACAATAATA	CCTAGTATCA	ATACAATACC	GTTTTGATGT	TCTAATAATT	2460
	TGGACCTTT	AACAGCAGCG	ATAATTCCTA	AATTACTCCA	CTTTAAATAC	GCAAGTAGCT	2520
	GTGCTGCAAA	AAACACAATA	ACGATAAATG	TTCCCATGTA	TCCTACAGCA	TCGCCGAACA	2580
20	TTTTACCTAA	GTCTTTTGTA	TTTTTAATTT	CTTTGCTTAA	AATCCCATAA	ACTAATCCAG	2640
	GTACTAAAAA	TACGACAAGA	ATAATTAATC	CGACACCGTT	AATTAATGGC	GCATCGTCTA	2700
	GTAAGCTGCC	TGTTTTAGCA	TTTCTTAAAA	AACATATGTT	AGGAATAGCT	GTAATAATTA	2760
	ATAAAATAAT	TGTGACTATG	AAACTGATAT	TTGCCCATTT	TAAAGCATGT	GCTTCTTTGT	2820
25	CAGTAATATG	TGAAGATGCT	TCTTCAGGGT	CATG			2854

## (2) INFORMATION FOR SEQ ID NO:67:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1442 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

	CTATGATAGN	ATCGGAAAGA	TGTATAAAGT	TATCTAAAAG	TTATACGACA	CAAGTACACG	60
	AGGCACATCG	CTATGCGGTG	TGTCTTNTGG	TATGCAATCA	AAGAGGTGTA	AGAGATGACC	120
45	AAGCATAATA	ACATTIATTA	GCATGGNCGT	AAGTCATATC	AATACGATTG	GTTCTATCAT	180
	TCAAAAGCAT	GGAAGAAGTT	AAGAGAGATA	GCATTAGATA	GGGATAATCA	TCTTTGTCAA	240
	ATGTGTTTAC	GTGAAGACAT	NGTAACAGAT	GCAAACATAG	TGCATCATAT	TATTTATGTT	300
	GATGAAGATT	TTAATAAAGC	TTTAGACTTA	GATAATTTGA	TGTCAGTTTG	TTATAGCTGT	360
50	CATAACAAAA	TTCATGCAAA	TGATAATGAC	AAAAGTAATC	TTAAGAAAAT	TAGAGTATTA	420
	AAAATTTAAA	TAAAAAATA	ATTTATTTTT	ATAGCCCCCT	ACCCATCGGC	TTAAATGTT	480
	TTTTCGACGG	GTACCGGCGG	GGGCCCTTCG	CTTGCAACGC	GGATAAACTT	TTATGAAAGG	540
	GGGTCTTTAT	ATGAAATTAA	CAAAAAAACA	GCTGAAAGAA	TATATAGAGG	ATTATAAAAA	600
55	ATCTGATGAC	ATATTAATTA	ATTTGTATAT	AGAAACGTAT	GAATTTTATT	GTCGGTTAAG	660
	AGATGAACTT	AAAAATAGTG	ATTTGGATGA	TAGAGCATAC	AAACAAGGCT	GGTGCCGAGC	720

EP 0 841 394 A2

	AATATTGT	TGA	AGAATCCATT	AAGCATAGAA	CTGACAAAAA	CAGNTCAAAC	ACTAAATAAC	780	
	TTACTCAA	AGT	CTATGGGTTT	ANCTGCAGCA	CAAAGAAAAA	AGATAGCNCA	AGAAGAAGGT	840	
5	GGATT	CGGTG	ACTATTAAAG	TTTTAAATGA	ACCTTCACCA	AAACTATTAA	CAACATGGTA	900	
	TGCAGAG	CAA	GTCAC	TCAAG	GGAAAAATAAA	AACAAGCAAA	TATGTTAAAA	AAGAATGTGA	960
	GAGACAC	CTT	AGATATCTAG	AAAATGGAGG	TAAATGGGTA	TTTGATGAAG	AATTAGCGCA	1020	
	CCGCCCT	ATT	CGATTCATAG	AAAAGTTTTG	TAAACCN	TCC	AAAGGATCTA	AACGTCAACT	1080
10	TGCATT	TACAA	CCATGGCAAC	ATTTTATTAT	TGGCAGTTTG	TTTGGTTGGG	TTCATAAAGA	1140	
	AACAAA	ACTG	CGCAGGTTTA	AAGAAGCTTT	GATATTTATG	GGGGCGAAAA	AATGGTAAAA	1200	
	CAACTAC	TAT	ATCTGGTGTT	GCTAACTATG	CTGTTCTCA	AGATGGAGAA	AACGGCGCTG	1260	
	AAATCC	ATCT	NTTAGCAAAC	GTAATGAAAC	AAGCTAGGAT	TCTATTTGAT	GAATCTAAGG	1320	
15	CGATGAT	TAA	AGCTAGCCCA	NAGCTTAGAG	AAAAATTTAG	ACCTTTGAGA	GATGAAATTC	1380	
	ATTACG	ATGC	ACTATATCTT	AAAATATGCA	CAGGCTTCAG	ACAGTGATAA	GTGGTTGGTT	1440	
	AA							1442	

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 332 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

35	GAATTCAAGC	CGACCGATGA	NGGTTATNTG	ANATTACATA	AGACGTGGTT	TNGTNANTCA	60
	AAGCTATGTC	CAGTTTGTA	TTCGGAGGCGT	GCTATGAAAA	NTAGTTATCA	CGCTCANANA	120
	GTGATTGANG	ANGTAATTAA	GGAAAAGCCA	NCAGCACGTT	GGTTGTTTTT	ACCACTNTCA	180
	CCAANANNTG	CGATAGATGG	AGATACTTTA	GTAACAAAGT	TNGANGCATC	TAACTAANGC	240
40	ATTTGATAGG	TTGAGTAGCA	TATNAAAAGG	TTAAACANAA	TCTNGTTGGA	TCTATGCGTN	300
	CAACAGAAGT	TNCCGCCTAC	CTAAAAATGA	CG			332

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1070 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

EP 0 841 394 A2

	CCATCCGTTT	NCNTCATCNC	ANGACNTTCN	ANANTTAATT	TTTACNAGGA	GTGACATCTG	60
5	TGAATAACAC	ACAATCCTCA	CCACGCAGTN	ATANTATTAT	TGCGATTATG	TTGTCTGCAT	120
	TAACATATTG	GGTGTGTGNA	CAATCATTTA	TTAATATAGG	ACCTCTCGNN	GGNCAAACGT	180
	ATCAAACCTC	TCCTGCCGTG	TTAAATTTAT	CTATTAGTTT	AACTTCCCTC	GCCACAGGTA	240
	TCTTCATGGT	GGCTGCAGGT	GATATTGCTG	ATAAAATAGG	ACAANTGAGA	ATGACATACA	300
10	TGGGTCTCAT	AATCAGNATG	TTGNATCTC	TTCTATTAAT	TATATCGGAC	ATCACTGCAC	360
	TGCTCATCAT	CGGNAGGAAT	TTTACAAGGT	CTATCAGCAG	NTATCTTGGT	TACCCTCCAA	420
	CAGTTGGCGT	TTTAAATAAT	CAATTTAAAG	GAGAACATTT	AAGACGAGCG	ATTAGTTATC	480
	TAATGATTAG	TACTGTTGGG	GGCATCGGCC	TAGCTGGTGT	TATCGGCGGT	TTAATTGCTA	540
15	CAAAATTCGG	ATGGCAAATG	AATTTTCATCA	TTAGTATAGT	CATTGNTTTC	ATTGCCATAT	600
	TACTTCTAAA	AGGCACACCT	GAAAAAGTAA	GTCAACATAG	NCACCGTCAT	CCGTTCGATT	660
	ACAAAGGTAT	GTCGATTTTC	GCTGTTATGA	TTGGTAGCTT	TACATTATTG	TTAACACAAG	720
	GATTGGAACA	AGGTTGGTTT	AGTACATTCT	CATTCAATTT	TCTGAGCATT	TTTATCATCA	780
20	CTACGCTGAT	ATTCATCATC	ATCGAACGTC	GACATGAAGT	ACCCTTTTAT	TGATTTCTCA	840
	GTATTACGCA	ACCGNCCGNT	CATTGGTGCA	TTTTTAAATA	ACTTTGTTTT	AAATAGCGGT	900
	CTAGGCGTTA	CAGTGGTCTT	TTTCATATAT	GCTCAAACAC	ACCTTGGTTT	ATCAGCTGCG	960
	CAATCTGGAT	TGNTACATTG	GCATATNCCA	TAGTGGTAGT	TGCGATGATT	CGTCTTAGGT	1020
25	GAAAAAAGCA	ACATTACGGT	TCGGGTGGGC	AAATTGGATG	CTCATCATGG		1070

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1246 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

	GAACAANTGG	TAAAACCACG	ACTTCNAACT	TAATTGGACN	TACTTTAAAA	GCAAATAATA	60
	TNCNAATTAT	ACNCAATAAT	GAAGGTGCTA	ATATGGCTGC	AGGTATAACT	TCTGCATNCA	120
45	TCATGCAATC	AACACCTAAG	ACTAAAATTG	CGGTAATCGA	AATTGATGAA	GGTTCGANTC	180
	CACGTGTGTT	AANAGAAGTT	ACACCTNCAN	TGATGGTATT	TACTAATTTT	TTNAGAGATC	240
	AAATGGATCG	CTTCGGTGAA	ATTGATATTA	TGGTTAATNA	CATNGNAGAG	ACAATTAGTA	300
	ATANAGGCAT	CAANTTATTG	CTAAATGCTG	ATGATCCATG	TGNGAGTCGG	GTGAAAATCG	360
50	CAAGTGANAC	GATTGTGTAC	TATGGTATGA	AAGCACATGC	CCATGAATTT	GNACAAATGT	420
	ACCATGAATG	AAAGTAGATA	TTGNCCAANC	TGTGGTCGCT	TATTGCAATA	CGATTATATT	480
	CATTATAATC	AAATNGGTCA	TTATCACTGT	CAGNGTGGTT	TCANACGAGA	GCAAGCAAAA	540
	TATGAAATAT	CANGTTTTGA	TGTGGCACCG	TTTCTATATC	TANATATCAA	TGATGAAAAAN	600
55	TNTGATATGA	AAATTGCAGG	TGACTTGAAC	GCTTATAACG	CGTTAGNAGC	ATATACTGTT	660
	TNAAGAGAGC	TAGGGTTAAA	TGAACAANCA	AATTAANAAT	GGCTTTGAAT	ACGTATACAT	720

# EP 0 841 394 A2

CAGACAATGG TCGCTATGCA GTACTTTTAA NANAGAACGA AAAGAAGCGA TGA CTCAANT 780  
 TTAGCTAAAA ATCCTGCAGG AATGAATGCA NAGTCTATCA AGTGGGTGAA CAATTAGAAG 840  
 5 GCGAAAAAGT GTATGTTATT TCGCTAAATG ATAACGCTGC AGATGGTCGA GATACTTCAT 900  
 GGATTTATGA TGCAGATTTT GGAAAAATTA TCTAAGCAAC AAATTGAAGC TATCATCGTG 960  
 ACAGGTACAC GAGCAGAAGA ACTTCAATTG CGATTGAAGT TAGCAGAGGT TGAAGTACCA 1020  
 ATTATTGTTG AGCGTGATAT TTATAAAGCA ACGGCAAAGA CTATGGATTA TAAGGTTTCA 1080  
 10 CAGTTGCAAT ACCAAACTAT ACATCAATTA GCGCTATGC TTGACAATTA AACCCTCCGT 1140  
 TTGAAGGAGG GCAATCATAA TATGCATGAA TTGACTATTT ATCATTNNAT GTCAGATAAA 1200  
 TTGAATTTAT NCAGTGATAT AGGAAATATT ATTGCTTTAA GACAAC 1246

## (2) INFORMATION FOR SEQ ID NO:71:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 751 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

CATTCAAGAA TTATCTATAC TTTAAGCCCG ACTAAAATCT ATAGAACAGC ATTACTAAAA 60  
 GATAATGACA TTATTTTCC AGAAGAATTA AAGAGTGCAG AAGATCAATT ATTTACAATG 120  
 AAAGCATATT TGAATGCAAA TCGAATCAGT GTGTTAAGTG ATAAAGCGTA TTATTATGCT 180  
 ACAAAGCGTG AAGGTGAACA TATGAGTAGT GCGTATGTTT CACCTGAAGA CTTTTACGAA 240  
 35 GTCATGAGAT TGATTGCTGT AGAAATATTA AATGCAGATT TAGAAGAAGC TCATAAAGGA 300  
 TCAAATCTTA GCAGGAATTT TTAAATCCGN CATTTTAGTT TTTCTCCGTA CGAATGGCTT 360  
 CTCACTTAAA GTTAAACTTG AAGAGCAACC ACAATGGATT AATGCTCTAG GAGGACTTTA 420  
 TACAAGCAGT TCCAGAACGT GTAGATGCAT TGGTGATGAG TAAATTACGA CCATTGTTGC 480  
 40 ACTACGCGAG AGCGAAAGAT ATAGACAACT ATAGAACTGT AGAAGAAAGT TACCGTCAAG 540  
 GTCAATACTA CCGTTNNGNT ATTGTAGATG GTAAATTAAA CATTCAATTC AATGAAGGCG 600  
 AACCATACTT TGGAAGGCAT TGATATCGCT AAGCCAAAAG TGAAAATGAC AGCATTTAAA 660  
 TTTGGATAAT CATAAAATTG NTACAGGAGC TAACGGTTAA ATGATTTATG ATTGGCGAAG 720  
 45 GGACAATTAT GATGTTCAAA GCTTTAAATT T 751

## (2) INFORMATION FOR SEQ ID NO:72:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1394 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

# EP 0 841 394 A2

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

5

10

15

20

25

30

35

```

CCTGTTTGCC CTAATGGANT CGGGTNACAA TNCACGGAGA TCCTTCTTCT CACATCAATN      60
TNAATG GGGG TGTTCAACCT CATCTTTGAT CTGATGACGA TGACATGCGC GGCATGCAAC      120
TTGTCCGNTC AATTTGTGTC AGATCATTCG TACGNATAAA ATTAGTAATN TTAACGTGTA      180
TCGGNAATNT TAAATAAATG TTTAATGCTT TTGAGAATAC AGTTCTANCC GNAGACGTCG      240
GGGACACTAA ATCACCGGAG ACGCCAAGGT CTTTTATTAA ACCTTGTTCA CTATATTGCA      300
TATACTGTGG ATACTGTCNC AACACATTGA NTTGATAAGG ATGTGTTGGT AATAAAATAA      360
AATCTTTGGG TATCTCTGAT ATATCTATGT CTGCTAATTG ATACAACACT TTCTCAACCT      420
GATCTTCTTT ACCTTCTACA TAGCGCGTGA GCAGAACATC TGGATGCACA GCTAAATAAT      480
GCAATTGGAA TGATGTATGA CATTTCGGGTG CATATTTCTC TAAATCTGCT TCTGAAAACC      540
CACTTGCACCT CTTAGGAGTC GGGATGAAAT GGATGACCTA AGTATTNAAG ANTGGTCTGA      600
AACGATATAA CGATCCTCTA CGTAGGCTAT TGTGTTACTT GGGAAAGAAC GCGCCGNGCG      660
ATGAATGCTA TTATCGATGN CAAACATAAT TNGGCCATA TGTTGGTTGC ACTGCCGTTT      720
GATTATCTGC ACTTTGAGCC ATATGTGGCA AAATACGCGC AATTGCTTCT TTATAAGTTG      780
TTATTTTTTT ACTTTTCCCA TCGATAAGCC ATACCTCTGG ATGATACATA TGATGCCCCA      840
TCGCAGACCA ATAGCGAAAT TCACCCGTTA AAGTTTCGAG CTCTGATAAT TGTATAGACC      900
ATTGATGATT TTGAGGTGGT ACTTGATATA AATTTTCTTC TCTAAAATAT TCATTTAAAA      960
TGCGTTTCGAT AGCCGCATAC GCTGCATGTT GTATTAATTC TTTATTTTGC ACTTTTTTGGT     1020
TTCAACTCCC ATAATTTTCAT TAATGTGTGA TCGGTTGATT TGATTAGTGA TGGTTGAACA     1080
AATTA AAAAT AAAC TACTTA CTGCAAATAC TACGCCATA ACGATAAACG TAGTAGCTGG     1140
TG TAGTATAA CTTGTAATGG CAGCGCCACT AAGACTGCCA ATAATTGAC CAACA ACTAA     1200
CATACTGTTC GTCGTTCCAA CAAATGTGCC NNTAAGNTGT TGATGACACG CANTCACGAC     1260
AACAAACATG ACAC TTTGAA TCNATGCACT AGATGTTAAT CCTTGAAGTA TTCTTG CAGC     1320
CATTAANAC TCTATATTCG TCGCTAAACC TTGCAGTATC GCACTACAAC CACATGCAAT     1380
CGTGGCAAAT ATAG                                                                1394

```

(2) INFORMATION FOR SEQ ID NO:73:

40

45

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 459 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

55

```

ATAATATGTC TAATTGACTA ACTTGTTGGA GTCATTTACT ATTTTATGTA TGACATATTT      60
TAAAAAGTGA GGGTCAAGCA TGTCTTATAA AGCATATCCA CTCTTTAGAG ATATATTAAT      120
AAATGAATGT ATTTATTTTCG CCTCTAAAAA TAAAAAACTA GTACGCCTAA ATTATAAAAG      180

```

# EP 0 841 394 A2

TGAAGCGTAT	GTAGGCGTTT	GGACAGAAGA	AAGTGTGGCC	GTATCATTTT	TAACAAGTCG	240
TGATATTCCA	TNTGATAAAG	TTGTAAAAAT	GGACGTTGAT	CCGCCGNGCT	ACTTATGAAT	300
TAGATGAATT	GTGTGATGAA	CAAGACATAT	TATTATGAAT	CAAACAATGG	AAGAAGAAGG	360
GCATCTACTA	ACGTGGGGCT	TGNTACAAAA	AGAAGGTGAT	GACGGGATTA	GATAAAANAG	420
ATCAAAGATT	TGGGCCCCAGA	TGTTGCAAAT	ATGATGAAG			459

## (2) INFORMATION FOR SEQ ID NO:74:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 930 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

CTGGTGCGTC	TTCTAATGAA	TTTAGAGATT	TTTTACCAAA	TCACCATATG	CAGTATACGA	60
TGATGAAGTA	TGCACGTGAA	CATGGTGCAA	CAACTTACGA	TTTCGGTGGT	ACAGATAATG	120
ATCCAGATAA	AGACTCTGAA	CATTATGGCT	TATGGGCATT	TAAAAAAGTG	TGGGGAACAT	180
ACTTAAGTGA	AAAGATTGGT	GAATTTGATT	ATCTATTGAA	TCAGCCATTG	TACCAATTAA	240
TTGAGCAAGT	TAAACCGCGT	TTAACAAAAG	CTAAAATTAA	AATATCTCGT	AAATTAAAC	300
GAAAATAGAT	TAACGACTGA	AATCTGAACG	CTCATAAGAC	TGTCATTTGC	GTTTCAGATT	360
TTTACACAA	TATAGAATGG	TTGAGTAAAA	TATTTTGTAA	TATAGTGAAG	GAGGGGGAAG	420
TACTGTGATA	AAAAAGCTAT	TACAATTTTC	TTTGGGGAAT	AAGTTTGCTA	TCTTTTAAAT	480
GGTTGTTTAA	GTTGGCTTGG	GCGGTGTATA	CGAGCGAAGT	GCTAAATTGA	AATTAGAATT	540
ACTACCAAAN	TGTACAAAAA	TCCAGTTATT	TCAAGTTACA	ACAACAATGC	CGGGGTGCAA	600
CGCCACAAAG	TACCCAAGAT	GAAATAAGTA	GTAATAATTGA	CAATCAAGTA	AGATCGTTGG	660
CATATGTGAA	AAATGTTAAA	ACGCAATCCA	TACAAAATGC	TTCAATTGTA	ACAGTTGAAT	720
ATGAAAATAA	TACAGATATG	GATAAAGCAG	AAGAACAGCT	TAAAAAAGAA	ATCGATAAAA	780
TTAAATTTAA	AGATGAAGTT	GGTCAACCAG	AATTAAGACG	TAATTCGATG	GACGCTTTTC	840
CGGTTTTAGC	ATATTCATTT	CAAATAAAGA	GATGACTTGA	AAAAGTACGA	AAGTACTGAA	900
TGACAATTAA	TACCAAATAC	AAACGGAGAT				930

## (2) INFORMATION FOR SEQ ID NO:75:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 744 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

# EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

5	TTTACCTTCC TTAATTGCTT CCAACCCAAT CATCAGCTAA TAATCCACTA CTTTTTCACC	60
	CATGCTCTTT CACNAATGTT TTGAACTGCG TTTCCCTTTAA CGTCATACCA ACCTGTTTTA	120
	CCTACTTTTG TAATATTACT TGCCAACACC ATTTTATTAT TATAAGCGAT TTGGTTAGCT	180
10	ATCGTGAACA TAGGTTTTAA ACATCGTCTG TGCACCCATA AAGGTATGCC TATCCAGGTT	240
	TTTTGATTAC TATCAGTAGC ATCAGATTTT CAAAAACCAT ATTGATTGGC GTTGTCTGCA	300
	ACAGACTGCA CAGAAGCTTC TTTAGAACT AGATATTCCG GAACATGATA ATTTTACGA	360
	ATGTTATCAA TTAAATGACT TTCTAAAGTC ACAACCGGTT CTATTTGAAT CGGATCACCT	420
15	ACAGCTACAA CTTTTTTTGA ACGATATAAT GCTCCACAG CTGCTTGAGG TATTGCTTGT	480
	CCTGCTTCAT CAATAAATAA GTAGTCTATG AAATCTTGTG GTATGCCCCC ATACATAGAT	540
	TTAAAGCTTG CAAACGTCGN ACTAACTACT GGAATATNA AATGCATCAC ATTCCATGCG	600
	TNGTGTNCTT TATCTGGATG NGCATCAATT AATTTCTTC TATCTNAAA ATCATTAATC	660
20	GCATAATAAA TAGTTGGTTT TATTAGCAAT CAATAATAAT TTATGCAATA TCAATNGCTT	720
	CTTAAAAAGA GCATGGACCG TCTT	744

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 858 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

	GGCCTGGATA AGTAGTACCA GTCCACATAA TCTAACCTAA GACAATTAAG ATTAAGATTG	60
40	ATCGCATGAT TCAATTTATT TAGCTTTGTA ACCCAATCAT ATTGANTAAA TCTTTAGGGT	120
	GCCAATATGG TGGTGCATAA NCCACNNCAA ACTCAGTTAA CTCATCTACA GTTAGCTGGT	180
	TCATCATTGC CATCGATAGT ACATCAATAC GTTTATCTGC ACCTTCTTTT CCTACTGCAG	240
	CTGCTCTTAA AATCTGACGG TTTGAAGTGT CATAATATAC CCTTAAGTGT AAAGGGGAAT	300
45	TTCTTGGGTA ATAATTCGCG TGTGCACCTT GAGTGACTTC CACCATTTTA TAGTCAAATT	360
	GCTTTAGTTC ATTTGGTTTA ACGCCGACAC TCGCAAATGT ATAATCAAAG AACTTCACAA	420
	TATTGTTGCC TAAGAAGCCT TTGAATTCAT TAGTGTCAAT TCCAGCAATT TGTTCGGCAA	480
	CAATACTTGC TGCACGGTGA GCGCCCCAAG CTAAAGGAAC ACTAGCCGGT AGATCGACAT	540
50	GTCGATAATG TGATGTTGCA ATATCGCCTA TTGCATAAAT GTTTGGAACA TTTGTTTCAA	600
	ATTTATCGTT TACGGTATGA AACCTTTTCG ATCAAGTTTG ATATTTGAAC TTTTCGATAAA	660
	TTTTGAATTG GGGTGAGTAC CGACACCTTC AATAATCATA TCGTAATGTT CAACTTTTCC	720
	TGATTTAAAT GTAATTTTCAT TTCCATTGAT AGCATCAATT CCTCATTTAA CGGTATGGAT	780
55	CTCCGCTATC TATTCAACAA GTATTGGTTG CTTCATGTCA GCATCCATAT TTATTTATCT	840
	TATCAGATCG ATGATTAA	858



## (2) INFORMATION FOR SEQ ID NO:77:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1004 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

```

GGAATGCGAT ATGCCTTTTT ATTCATCGTC ATATAAATTG CATTTTCATT TTTACGTGTC      60
ATACGCTGAA AAGGATGCTT CACAAATTGT GGTGTAGTAT AACCAATTCC AATTTCCTAAA    120
TACAACACTT TATCATCTTG ATGTTGCTCT AGAAAAGCAT TATAACGTTG NNGNTGNGCC      180
TGAAATTCAG CATCTTCAAC CATGCCAACT TCCGCTTTAC GTTTATTCAC TTCCATTGGA      240
GCATCACATT NTGGACATCT TGGAATCATC TCCAAGGTA TAAGCATATC TTGTTGCGCA      300
ACAACCATTT TACGAATTAA ATCATCATTG CGATACGTGT GAGCATGACA ATGCTGACTA      360
CACTGNTGNA GTATATACTC CCCNTGTATA TGAAATACAT GAGTCATATC ATATTCAGCA      420
GCATCGAAAG CATTGTCTGC ATTCGGAGNT ATGATATGGT ACTGTTTACC CTCCACTAAG      480
GATTTTAATG CGAGATAAGA CTGACCTACA GGTGGATCT TAATAATTTA NTGTAATAAA      540
ACGACTCTCA AATGCCCAAT ACTCTTGCCA ACTGCCATAA GGATGTAAAC TCGCTTGCAA      600
CATATCAAAG AAGCGATATT TTTCAATAAA ATCTGGGAAA TTTTCGTAA AACGCTCTCC      660
TACATATGTA AATCCGTCAG ATGCAGACAT GCCTGCACCA ATTCCAATCA CTATCGCATC      720
TGCTTCATCA ATCGCAGTAC GCAATACTTC AGCCTGCTTT GTCTTTTCAT CCATTAAAAG      780
AGACATTGCA TTCCATTTAC TACTCTGCAT CACGGCTCAA TGCCTCCTTA TACAGTTGGT      840
AAATCCTTAT CTTGGAAATA CATTGAACAC GACTTTCAAT GTTGAATTTG GCTCTGCGAG      900
ATAGCTTTCT TCTGTTCGAC AGCATTTCGT CTGCTTCATC TTGAGAAAAG CAATACACTG      960
TAGATATACA GCAAAGCGAC ATGATTTAAC TATGTGGGCC AGGC                      1004

```

## (2) INFORMATION FOR SEQ ID NO:78:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 717 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

```

ATATTAGCAG CTGTGCTTGC TTGTATTTTA GCTGGTGAC CCTATCAAAG CAATTATTGA      60

```

EP 0 841 394 A2

	TAGTTTAACT	ACATTTTCAG	GAATAGAGCA	TAGATTGCAA	TATGTTGGTA	CTAATAGAAC	120
	TAATAAATAT	TATAATGATT	CCNAAGCAAC	AAACACGCTA	GCAACACAGT	TTGCCTTAAA	180
5	TTCATTTAAT	CAACCAATCA	TTTGTTTATT	GTGGTGGTAT	TGGATCGAGG	GAGATGAATT	240
	TGACGAACTC	ATTCTTTATA	TGGAAAATGT	TNCGCACGAT	GGTTGTATTC	GGACAAACGA	300
	AAGCTNAGTT	TGCTAAACTA	GGTAATAGTC	AAGGGAAATC	GGTCANTGAA	GCGAACAATG	360
	TCGAAGACGC	TGTTGATAAA	GTACAAGATA	TTATAGANCC	AAATGATGTT	GTATTATTGT	420
10	CACCTGCTTG	TGCGAGTTGG	GATCAATATA	GTACTTTTGA	AGAGCGTGGA	GAGAAATTTA	480
	TTGAAAGATT	CCGTNCCCAT	TTACCATCTT	ATTAAAGGGT	GTGAGTATTG	ATGGATGATG	540
	AAACGAAGAN	CGATCAACAA	GAATCAAATG	AAGATAAAGA	TGAATTAGAA	TTATTTACGA	600
	GGAATACATC	TAAGAAAAGN	CGGCAAAGGA	GAAAGGTCAA	AGGCTACACA	TTTTTCTACN	660
15	TCAAAATAAA	GATGATACAT	CTCAACAAGC	TGATTTTGAT	GAAGAAATTT	ACTTGAG	717

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 650 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

	CTCTTCTTCA	TCAACATCGA	TCATACTTAC	TAAATTGTGG	ATGTGAAAGC	TGTGATGAGT	60
	ATATGTACTT	CTCGMNCAAA	ACGGTTTAAT	GTNTCCTCCT	TTTCTCCAGG	GTGGTATAAA	120
35	AATCGCCTTA	ATTGCAACCT	TAATGTTAAG	TATCGTATCT	TCTGCAAGAT	AAACGGTACT	180
	CATGCCACCG	NCGCCAAGCT	TATCTACAAT	TTTATATCGC	ACATTTATTA	TTTTACCTAT	240
	CATACTTTAT	CACCTNCAAT	AGCCGCGAGT	ATGAAAGTAA	CGTTATCTTT	CGAATGGTTA	300
	TCTAATGCCA	ATTGCATTAA	TTGATCACCA	TGATCTTCTA	TTGTACCTTC	TTTTACTAAC	360
40	AAACGCTTAA	TTTCATTGTC	TTTAACATAA	TCAGTTAATC	CATCTGAATT	TAATAATAAA	420
	TAATCATAAA	AATTTAATCG	CTTAATAAAC	AAATCTGGAC	TCACACGTTT	ATCTGTGCCC	480
	ATCACCTTCG	TAATAATATT	ACGTTGTGGA	TGTGTAAATG	CTTCTTCCGG	CGTAATTTGA	540
	CCCGTTAAAA	CAAGATGATT	ACAAATGAGT	GATCACTAGT	AATTTGCACA	ATTGTCTACT	600
45	ATTAATACAT	AGGCTCTAGA	ATCACCGACA	TTTGGTATCA	CAACTGATTT		650

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 524 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

# EP 0 841 394 A2

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

```

CACCAAGAAC TTACTTATAT TTTAGACAAC ATAAAAGGGA ATAATAATTA TGGTAAGGAA    60
TTTGTGTCAG CCGNTGAAGA AACATTCGAC ATTGAATAAA GCGGGGTGAA GCACTATGAA    120
TCAATGGGAT CAGNTCTTAA CACCTTATAA GCAAGCGGTT GATGAGTTGA AAGTGAAACT    180
TAAAGGCATG CGCAAACAAT ATGAAGTTGG TGAACAAGCG TCGCCAATAG AATTGTGTAC    240
TGGTCGTGTT AACCCGCATC GCTAGTATTA TAGATAAGGC AAACAAACGA CAAATACCAT    300
TTGGATAGGT TAAGAGAAGA AATGTACGAT ATCGCTGGGT TTAAGAATGA TGTGCCAATT    360
GGTAGAAGAT ATTGATGTTG GCCGCCAATA TTTTAAAGGA CAAAGAAAAG ATTTTAAAGN    420
TATTGGAGAC CGAGATTATT TCCCGNACAC TAAAGGAAGG TGGGTACCGT CCGCTTCANG    480
TCAATATTGG ATTTCCCAAC TGGACCAATA CAAGGCCAAA ATTT                    524

```

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 751 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

```

CGTCTATCAT TGTTAACATA TATAGCACCT TCCTTATTTT AATGTTGTTT TAGTTGAATG    60
ACAGTAAAAA GGTGTGTTAAG ATACTCATAC ATTTTATGTT GTAAATATCT ACAAAGTTAC    120
CCANCTACTG ACAATGTTTA TTTNAGATAG TATATGTAAA TTCACAGATA TGCTAATTGC    180
TTAAAAAATG ATTAAAGTGT TGGCTCCAAG CAATGATACT TTAGAAATTT ATTTATCATC    240
TNGACTTTAA AAATTATATT ATAAATGACG TAACTGACAA CAGATATACT TAGTAATGAA    300
GATGTGTAAT GTAATTGTTT AAAATTGATC TCCAAGCAGA TTTTATTTAT CATTTAATTT    360
AAATAGCAAG TGGAGGTACA AGTAATGAAA TTTGGAAAAA CAATCGCAGT AGTATTAGCA    420
TCTAGTGTCT TGCTTGNAGG ATGTACTACG GATAAAAAAG AAATTAAGGC ATATTTAAAG    480
CAAGTGGATA AAATTAAAGA CGATGAAGAA CCAATTAANA CTGTTGGNAA GAAAATTGCT    540
GAATTAGATG AGAAAAAGAA AAAATTAACT GAAGATGTCA ATAGTAAAGA TACAGCAGGT    600
TCGCGGTNAA AGCAGTAAAA GATTTAATTA AAAATGNCGA TGATCCGTCT TAAAGGAATT    660
TGAAAAAGAA GAAGACGCAA TTAAGAANGT CTTGAACAAA GACTTTAAGA AAGCAAAAANG    720
TACGTNGGGA TAACATGATA TGATGTTAAA C                                751

```

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 785 base pairs

# EP 0 841 394 A2

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

CCTTTGTTTA	NCATCCTTAG	TCAGTGCAT	AAGGTTGTCC	CATCAGAAGG	TAGNCCCCAC	60
AATAGTTGAA	ATCCTCCAC	AATAGGGGTA	GNCTCCTGCA	TCTGTAAATA	CCTACTGGNN	120
TGCCANATAC	TTTTTCTGCT	GGTGTGTCAA	AGAAATTGT	TGTTCCATTT	GAGAGAACAC	180
TAATAATTTT	GACATAATCT	GCATATCGCA	TATAAATTGT	TGCGTTATCA	CGATAATCTT	240
CATGTAAATC	TGCTAAAGCG	TTAATAATAG	CATCATACAT	GTCTGCTCCC	NCAACTTCTT	300
TAACAGATCC	ATTATAAAAT	GACATGTGTT	CTAATCCAGA	TTTAGGACTT	ACTGCTAAGG	360
CATCTTTACG	CTCTTTAGCT	GCTAATCCTG	ATTGTAGTGC	GTTTTCAACC	CAGTTTACTA	420
AATCTACATC	TGATCCATGA	ATTACAGTAT	CTGAAATTGC	AGCAAATACT	TTGAANTTAT	480
TAAGTAGTGA	ACTTGACTGT	ATCACCTTGT	GCTTTTAATT	CTTGTGCTGT	NTCTACGTCT	540
GTAATGAAAT	CATCATCGTC	TAAAGTGTAT	GAAACTCTTG	GAATCTCTAA	ACCTTTAATG	600
TTAGNTAGAC	GAGCTTTTAC	ACGTAATTGG	GTTTNNAGCA	AATGGCTCTG	AAACAATTTT	660
TTGTAGAAAG	TGGGTTGGGG	AAGAGCTTAT	CTCCACCTGA	ATCATTTCCT	GTTGGTAAAG	720
CGTGTAAATA	ACGTTGTGCC	TCCATTGAAG	GTTTTTCAAA	TTCATTGGT	AAAATCGCTC	780
GTGCC						785

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 924 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

CTTNATAAAN	ANCNATCNTA	TGCCAATTTG	ACCATTGTTT	GAATCCATAT	ANAAACCGGC	60
NACGGTTCTT	TTCAAATATA	ATAGTAAGTG	TATAATGAAA	ATGTAAATAT	TATTAAANAT	120
GGGGGTTTAC	TCAATGAAAA	TGAAACGTTT	TATAGCTATT	GTAATGGCAT	TATTTTTAGT	180
ATTAGNTGGT	TGCTCTAATT	CTAACGATAA	TAATGAAAGT	AAAAAAGATG	ACGCAGACAA	240
TGGTAAGAAA	CAAGAGATTC	AAGTTGCAGC	GGCAGCAAGT	TTAACAGATG	TAACCAAGAA	300
ACGAGCTTCA	GAATTTAAAA	AAGAGCATAA	AAATGCTGAT	ATTAAATTTA	ACTATGGTGG	360
ATCAGGGGCA	TNAAGAAAAC	AAATTGANTC	AGGCGCACCN	CTTGTTGACG	TATTNATGNC	420
TNCCNAANTN	CTAAAGATGT	AGATGCATTN	NNAAGACAAG	GAATNNAGCG	CATTTGATAT	480
CATATNAATA	TGCGNNNNAT	AGTCTAGTAT	TAATTGGTGA	TAAAAGATTC	AAATTACACT	540

# EP 0 841 394 A2

	TCAGTAAAAA	GACTTAAAAG	NCAATGATAA	ATTAGCATT	GGTGAAGTGA	AAACTGTACC	600
	AGCAGGAAAA	TATGCGAAAC	AGTATTTAGA	TAACAATAAC	TTATTTAAAG	AAGTCGAAAG	660
5	TAAAATCGTT	TATGCTAAAAG	ATGTAAAACA	AGTATTAAAT	TATGTTTGAA	AAGGGTTAAT	720
	GCGAAACAAG	GTTTTGTGTA	TAAAACTGAC	TTATATAANC	AANNCNNAAA	AATTGATACT	780
	GTAAAAGTAA	TTAAAGAAGT	AGAACTTAAG	AAACCAATCA	CATACGAAGC	TGGTGCTACA	840
	TCAGATAGTA	AATTANCAAA	AGAGTGGATG	GATTCTTAAA	TCAGATAAGC	TAAGAATATT	900
10	AAAGATACAC	TTTGCAGCAT	AAGA				924

## (2) INFORMATION FOR SEQ ID NO:84:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 536 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

	GGCACGAGCT	CGTGCCGAAC	TTACGGCACG	ACGCGATACN	ATAAATACAT	TTGTTACAGT	60
	TACGCTACTA	ATCGTTTGTG	CTACATCTTT	AAAATCAACC	GCAATACCTT	TGTGTTTGTG	120
30	AAGTAACGTG	TTTAACTCTC	TCGTTTGTAG	TAATAAGCTC	ATGAATTTCT	CTCCTTGTGT	180
	ATATTTTAT	AGAATAAATG	CACCTAAATC	TTTATNTGTT	GAAATTGATT	TTAATTTATC	240
	ATCAACATAT	TGTGGGGTAA	TATCTACAAC	TGCATTCTGG	ATACTTGGTG	CTTCGAATGA	300
	TAAATCTTCT	AGCATCTTNT	CTAAAATTGT	ATGAAGTCGA	CGGTGCAACA	ATGTNGTCTT	360
35	GTATCTTGGA	TTCACTTGGA	TAAGCAATCT	CAAGCTTAAG	GCGANTAATT	GCTTCCATCG	420
	GGTAAAGTTT	ACAGGNAACT	TCTTCTTNTT	TGGGAGCAAT	GCTTCATATT	TGGNNTAATT	480
	AATGACAATG	TGGGGTCTGG	NCAAAATTCT	TACAAATCT	TCCTNCCCGT	TANCTA	536

## (2) INFORMATION FOR SEQ ID NO:85:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 802 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

55	CTACCTTCAT	GTTTTGTAAA	GTTGATTAAA	GCATTCATTG	CTTGTGTATC	ACTTGCAATT	60
	TCAAAAGTTA	CTCTACCAAT	GTCGTGGTAA	TAAGAATGTT	CAGGCCCAAT	ACCAGGATAA	120

EP 0 841 394 A2

	TCAAGTCCTG	CTGAAATAGA	ATGTGCTAGT	TGCACTTGCC	CACCTTCATC	TTGAATTAAA	180
	TACATTTTAG	TACCATGTAA	TACGCCAGGT	GATCCTTTGN	CAATTGCAAG	TGCATGTTTA	240
5	TCAGTATCCT	CGCCTTGACC	TGCGGNTTCA	ACACCGNATT	AATGCAACAT	CATCTTTAAT	300
	AAATGGATAA	AAGGCACGAG	ACCGATTGCA	TTTGAGCCAC	CACCGATACA	TGCTACAATT	360
	GCATCCGGAA	GTGACCTTC	TTTCTTCAAT	ATCTGTGATT	TNATTCTTT	ACCAATGACA	420
	CTCTGAAAAT	CTCTAACAAT	CGTTGGGAAT	GGGTCTGGAC	CTAATGCAGA	ACCTAATAAA	480
10	TAATGTGTAT	CATCTACATG	ACTTACCCAA	TATTGCAATG	CTTTATTAAC	TGCATCCGAT	540
	AAAGTCCCTT	GANCTTCTAC	AACTGCCACA	ACCTTNGCA	CCAAGTAATT	CCATTCTAAA	600
	TACATTAAGT	TGGTGTCTTT	TAATATCTTC	ACTTCCCATA	AAGACAACAA	GTTCCATATC	660
	AAATAATGGA	GGAACCGGAG	GACTAGCTAC	AACATGTTGA	CCCGCACCAG	TTTCAGNAAC	720
15	AAGCTTCTTC	TTGGCCATTC	TTTTTGCAAG	CAACGGTGAC	TAACGCATAT	AATTTATGGC	780
	GCCGTATGAT	TAGATCCTCT	CG				802

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 662 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

	CTATTCTTCT	ATGTCTTTGA	CAAGCGCGAA	TATTTTTCGT	CGTTGCTTGC	CGGTAATTTG	60
35	AAACGGATCT	ATGACGCTTA	CATCGACTTC	CACATCAAAT	CCGTTATCAA	GTAAGTAATGT	120
	TTCTTTTATTG	CCTAATTCAA	CACCCGAGAT	GACAACTGTT	GNTGTACCGN	CATCTTGAGT	180
	GATATAACTA	GTAATTATTG	GCATCTAATC	ATTCCAATCA	GAACGGGAGG	TCTGAAAAAT	240
	CTTCTTCACT	ATTGTCAAAC	GGATTATTGC	CAGTTTGAGC	TTGTCTTGT	TGTTGATAAT	300
40	TGTTGTTTIG	NTGTTGGTTG	TTATTCTTCG	GTTCTAAGAA	TTGAACGCTG	TCCGCTACTA	360
	CTTCTGTCAC	AAATACACGT	CGCCCTTCTT	TGTTATCGTA	ACTGCGTGAT	TGTAAACGTC	420
	CATCAACGCC	AGCCAATGAC	CCTTTGGATA	AATAATTATT	TACATTTTCT	GCTTGTTTTC	480
	TAAAAGTTAC	ACAGTTAATA	AAGTCTGCCT	CACGTTCTCC	TTGAGCGTTA	GTAAATGTTC	540
45	TGTTAACTGC	GATAGTGAAA	GTGGTNACAC	TCACACCATN	TGGGCGCTGT	TCTATATTCT	600
	GGATCTTTTT	GGGTAAGCGT	CCCACTTAAT	ACTGTTCTCG	TNNTAACATT	ATTGNTTTC	660
	CC						662

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 656 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

```

10 CTGGTAAAC AACATTAACA GCAGCAATCG CTA CTGTATT AGCAAAAAAT GGTGACTCAG      60
   TTGCACAATC ATATGACATG ATTGACAACG CTC CAGAAGA AAAAGAACGT GGTATCACAA      120
   TCAATACTTC TCACATTGAG TACCANACTG ACAAACGTCA CTANGCTCAC GTCGNCTGCC      180
   CAGNATCCCG CTGACTACGT TAAAAACATG ATCACTGGTG CTGCTCAAAT GGACGGCGGT      240
15 ATCTTAGTAG TATCTGCTGC TGACGGTCCA ATGCCACAAA CTCGTGAACA CATTCTTTTA      300
   TCACGTAACG TTGGTGTAAC AGCATTAGTA GTATTCTTAA ACAAAGTTGA CATGGTTGAC      360
   GATGAAGAAT TATTAGAATT AGTAGAATG GAAGTTCGTG ACTTATTAAG CGAATATGAC      420
   TTCCCAGGTG ACGATGTACC TGTAAATCGT GGTTCAGCAT TAAAAGCTTT AGAAGGCGAT      480
20 GCTCAATACG AAGAAAAAAT CTTAGAATTA ATGGAAGCTG TAGATACTTA CATTCCACTC      540
   CAGACGTGAT CCTGACAAAC CATCAAGATG CCAGTGAGAC GTATCTCAAC ACNGTCGGGG      600
   TCTGTTGGTA CAGGCCCGTT GACGTGGGCA ATCAAGTGGT GAGGAGGTGG GATCAC      656

```

25 (2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:

```

30 (A) LENGTH: 390 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

```

35 (ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

```

40 GAATGCCAC ANAATGATGC TGGCAATGAT TTCACTTAG TTATGAGCGG TAATGAATTT      60
   TGCGGTAATG CGACGATGTC ATATATACAT CATTTGCAGG AAAGTCATTT GCTTAAAGAC      120
   CAACAGTTTA AGGTGAAGGT ATCTGGCTGT TCGGATTTAG TGCAATGCGC AATTCATGAT      180
   TGCCAATACT ATGAAGTTCA AATGCCACAA GCCCATCGTG TTGTGCCAAC AACAAATTAAT      240
45 ATGGGTAATC ATTCATGGAA AGCATTAGAA ATTATTTATG AAACATATTG TACATTATGT      300
   GATTCCNAGC TAAACAAAGT NACAACCTGA AATTCAACAT TTGGNTGGAA GCATTGTGTC      360
   CGTTGANCAA CAAATGGAGT CACAAATATT

```

50 (2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:

```

55 (A) LENGTH: 423 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single

```

EP 0 841 394 A2

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

GGNACGAGCG	CATNATTAAA	AACATATCAG	GTTATGTATG	CGCATAAAAA	ACAGCCTTGA	60
AAAACCTTTA	TAAATCGGAT	TTTCAAAGCG	TGTTTTGAAA	GGTATTTAAT	TAAAACTAAG	120
CATGTTGATG	TAAATCATCA	AAGTTTGTTA	AACGTTGTTG	CCACTCATTA	TCACTAATAT	180
TATTTGCTTG	AACATAACGA	TTACGCTCAT	GTTTAGCACA	TTCATAAGAG	CATGCACCTA	240
AATATTTAGT	TTCGNNTTCT	TCAGAAACTA	ATATTTGTTT	ATTACATTCT	GGGTTAGCGC	300
AATTAATATT	AACGNTCACA	TGGGTTTGGC	ATCAAAACCC	AATCCTTACC	AATAATTGTT	360
TTTTCAACTT	GGGTTGATAT	CAACACTGAT	ACGGATCAAC	AANTTACAAT	ACAATTTNAC	420
CCG						423

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 891 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

TTGGTGCACC	ACCTCAAAAC	TTTGGTACCA	TTGATGCATT	GTACATTAAT	TGCGCCAAGT	60
TTTTGTTAAT	GCCACCAATG	ATTGAGGTAA	TAATTGTCCC	AAAAGCAGCA	GTTACAATTA	120
AATCTACATC	TAATTGAANC	AATTGTTCTA	ATTCTTCTGA	TCCACTTAAT	TTTTCAGGCT	180
GATATACAGG	TAAATCATAT	TTCAATTNCA	CTTTTTTAAC	TGGTGGTGGT	GTCATAACAC	240
GTTTACGTCC	AACAGGTCGA	TCTGGTTGCG	TTACGACTGC	AATGACATCA	TGTTCTGCAA	300
TAAGCATTTC	TAAAACAGTT	GTTGAAAAGG	CACGAGTACC	CAATAAATATT	ATTTTAGTCA	360
TTTATAAAAT	ATGCCTCCAC	TTCTTTATCT	GTTAAAATAC	GGTCCGCACG	TTCNGTAAAA	420
GGGAGACCGT	TCATTTGATC	TATAATATGC	AAAATCATTC	TTGCTACATC	TTCATGTGCA	480
GTTAGTTCAA	CTTTGTTCCC	ATTGACGTCA	TAACTTTCGA	CAACTATCAT	TTTACTTCTT	540
GTCACCTTCG	CGTAAACATC	TGGCAATGTA	ATTGAACCTT	CTAAGTCTGT	TATTGTTTCA	600
TTTGATTGAC	TAATAATTTT	CGGATTAACA	AGTTGGTAAT	AATCCTTCCA	TTTCCATATT	660
CAATAATTGG	CACTTGGAAC	TGACTTGATT	AATTTGAGGG	TGCACATAAG	CCAGCAGCTT	720
CTTGGTGCAT	ACATTGGTAT	CTTCTTAAAT	CTTGGTAATA	ATCTTTTTTA	CGAATCACAA	780
TTGTTTTAAC	TTGCTTGCGC	TTTTTCCGGT	TAAATAGGAT	GCGATGCTGG	GTAATAACTT	840
TTTTAATCGC	CAATACTCTT	ACTCCTCAAT	AAATCAATCA	ACTATATACC	G	891

(2) INFORMATION FOR SEQ ID NO:91:



## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 652 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

```

GGCACGAGCT ATAAATCCAA TGATGAATTG TAAAAGTGAA TAATTGAGAA AAAGGTTAAT      60
ATCAAATTTT GGTGTCATCA TTAATGTAAG TTCCTTGGCT AACGTTGAGA AAGTTGTTAA      120
GCCACCTAAA AAACCGGTGA CAAAGAACGC AGGGAACCAT GAGATTGAAA TTGATAGGCC      180
TATAGTTAAT CCAATTAAAA AACTACCAAC TAGATTTACT ATCAATGTTG CGATAGGTAA      240
CTTTGAAGTA AATTTATGAT TAAAATAATC AGTAATGGCA CTTCTAGCAA TTGCGCCAAA      300
ACCGCCGCCA ATCATGACTA AAATGATTGA TATCATGATA AACCACCACC TAGTTTTATA      360
CCGACGTAAC ATAACAAAAT CCCAAAGACA TAACTTGTTA CAGCATATAG TAGTAAAGTT      420
ATAAATTGTT GATGATCAAA CATATGTATT AATTCCTAAT TGAAATGTTG AAAAAGTCGG      480
CTAAAGCACA AGAAAACAGT CGTAATAGCT TTTTTTAGGG TCGGATGGTT TGAAAAAATG      540
CAATCGTTAA GCTGTTAGCA TCCCATTACA AAGGCACCAG TCAATGGTAT CAGTGTCCGA      600
TGGAACTCCG CAGTATCAGA AAGATGAGGT ACGTATAAGG CTAAGCACAC CG              652

```

## (2) INFORMATION FOR SEQ ID NO:92:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1142 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

```

GAAACCATCC CCNCGTGCTC GTCAGTATT GTTCCGCCAC CACCTTGTGT TCCACCACTT      60
GATTTGATCA GCAGGATTGA AAGTACCTCT NCCACAGTAA TACCCATTAT TAGTCCCACC      120
AAAACCTGTA CATTACCTGC TGATGCTCCT TTCACCCATG GGCTCGTATT ATNACGACGT      180
GTAAATGTCG TNACCACATT TCCATTACGT TTAATAACTA ATTTGTCAGC ATATGTCGTA      240
ACATTACCAG CATGAGTATT GACTGTTTGG TTCGCACCAG GTGCAATTGT AATCGCTCCT      300
GCCGCTGTTT CAGTGACAGT TGGTTTCGCT GGTTGAACAT CTTTACTAC AAATTTGCT      360
GGTAAAGATG TTGCAAAATG ATGTCCATTA TAGATGACAT CATATTTTGC ATTAACGACT      420
TGTGCAGTAT TTGGTTTATT CATTGCTGCC CAGTTTGCAT CGTTTGTACC CGTAGTATCA      480
CGATTCCATT TATACGTAAA TCCATCTGTT GGTAACCTG AAGCGTTTTG CATATGTGCA      540

```

EP 0 841 394 A2

TATCCTGATG CTTGCGTACC ACTTGCTAAA GTGCCACCAA CTGTTGTTGT ATAAGTAGTT 600  
 TGAGGGAATT CCAAATTGAT ATACGTTTAC AGTTACAGGA ACTCGTTTAG CAGCTGAAAT 660  
 5 ACCTGGATAT GTGACATCGA CATTTAAATG TTGAACGCCT GCTTGCTGGT TATTTGGTTG 720  
 TTGTCTATTT GCCCATGCTG CTGTAATACC ATTCGTATTA GTATTGGAT CAAATGTAAT 780  
 GTAATCAATA GCGTTTGAC CATGTGTCAA ATTTTGACCT TTCACATCAC GTGATGGCGC 840  
 CTTAGCATTG GCAACTGGAT AAACTTTGAC TGGAACTTCA ACATTACGCG TACCTTGACC 900  
 10 ACTAGGTAAT GTTACAACCG CAGTTTATG TGTGTTGGCG ACTGTATTCT TCCATGTATC 960  
 TGGACTATCA TGCCATGCGA CCGCTGGCCC ATGTGGNGGA TTTTGGAATA AATCGTCTTC 1020  
 ATGACCGAAA TCAAAACCGN CGCACCTTTA ATAAATACAG CGCTTCAGTA GTTGCTTGTA 1080  
 TTGTGGGGTT ACTGTACAAG AAGCACTATC ATTTGATCAC AGATCAATCT TGGTNNTNGA 1140  
 15 AC 1142

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 605 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 25 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

GGNANTNGAG GAANCAAANC NGCCCTTCNA CNCGGATTAA AACCTGTTNG AANCTGNNGA 60  
 AGNGTGGGNC NAANNCCCTG ANTGCAAGGT GCGANGGNCG NNTGCAAGNT GTNNAACNGC 120  
 35 NNGNGANCGN TTGNCNGTGC ANTGNNAAG ANTGGTGAAA ACCCNTGTGN TANATTGTGC 180  
 GNCCNNCTTG GTGANGTGN GTTGNGCNGN NTCTTCAGNA GTCGCANCTG CAGNGTGTCC 240  
 NATAAGCGCT NINTGNACGG TTGCTGGTGT TGCNNANTCA TCTATCGCAA CATCGNTAAT 300  
 TGTTGTATCT CCAGTAATAC CTTGAATATC AGCAACTGCT TGATCATTA TTTGCGTAAC 360  
 40 ATCATTAGTT GNTTGTGCAN TTAAGATATC TTGANACGCT TTTTCTTTAG CTTNTAAAAC 420  
 TAAATCTTTT GCTGCATTTT TCTCTTCAGT TGTAGCGCCA GTTGATTAT CAATTGCTTG 480  
 ATTTTGAGTT GTCACAGCTT GATCAACTTC ATTTTTCGCA TTCGATTAA CTGCTGNTGC 540  
 TGGTGTGTG CTTTGAATTA NAGAAGACTC AAGCTTGCAT GCCTGCAGTC GACTCTAGAG 600  
 45 GATCG 605

(2) INFORMATION FOR SEQ ID NO:94:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1778 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 55 (D) TOPOLOGY: linear

# EP 0 841 394 A2

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

```

AAATTCATAA TAAAGNCTTT TGNCATTAAA GTTNNAAGA AGATAAAGAA NCTGCTNNAA      60
ATTGACGTAT CNNAAGAAG TAAAAGAAGT TTAATANTC AAATAATAAA TCNACAACAC      120
AAGANATTTC TGAAGAACAA AAAGGTGAAT ATCAAAGAAA GTCAGAGGCA TTAAGAGAAA      180
GATTTATAAA CAGACAAAAA TCTAAAAATG AGTCTGTGGT TTCACTAATC GATGACGAAG      240
ACGACAACGN AAACGACAGG CAACTTGTGG TTTCTGCGCC ATCAAAGAAA CCAACAACAC      300
CGACTACATA TACTGAAACA ACGACTCAGG TANCAATGCC TACAGTTGAG CGTCAAACCTC      360
AGCAACAAAT CGTTTACAAA ACACCCAAAA CCATTAGCTG GATTAAATGG TGAAAGTCAT      420
GATTTACAAA CAACGCATCA ATCACCACAA ACTTCAAATC ATACGCATAA TAATGTTGTT      480
GAATTTGAAG AAACGTCTGC TTTACCTGGT AGAAAATCAG GATCACTGGT TGGTATAAGT      540
CAAATTGATT CTTCATCTCT AACTGAACGT GAGAAGCGTG TAATCAAGCG TGNACACGTT      600
AAGAGAAGCT CAAAAGTTA GTTGATAATT TATAAAGATA CACATAGTTA GAAAAGACCG      660
ATTAAATGCA CAACAAAAAG TAAATACCTT AAGTGAAGGT CATCAAAAAC CGTTTAATAA      720
CCCAATCAAT AAAGTANCCA TGCCAATAAT ATTAATGCAT GGCTGCAAAG CAAATAATGA      780
GTTTGTGCGT AAAATACCAA CATTTAAACT AGCAATAAAT AATATCCAAG TCATCATTTT      840
ATTGATGCAA TCTAGTATAG TCCACATTCT AAACAGGTGT GGACTATTAC TTTTTCCTACT      900
TTATATTACC GAAAAAATTA TTATGCTTAA CTATCAATAT CAATAATTAA TTTTAAGCTG      960
AAAAACAATA AAAATGTTAA GACAACGTTT ACTTCAAGTT AATTATTATA CTGAAAATTC     1020
TGGTATATAA TGCTGTTAGT GAATATAACA GGGAAATTAT ATTGTTTATA ATATTGAGTC     1080
TATATAAAGG AGAAATAACA GATGAAAAAG AAATTATTAG TTTTAACTAT GAGCACGCTA     1140
TTTGCTACAC AACTTATCAA TTCAAATCAC GCTAAAGCAT CAGTGACAGA GAGTGTGAC      1200
ACAAAATTTG TAGTTCAGAA ATCAGGAATT AATAAAATTA TTCCAGCTTA CGATGAATTT     1260
AAGAATTCGC CAAAAGTAAA TGTTAGTAAT TTAAGTGACA ATAAAAACTT TGTTAGTTCT     1320
GAAGACAAAT TGAATAAGAT TGTTAGATTCA TCGGCAGCTA GTAAAATTGT AGATAAAAAC     1380
TTTGCCGTAC CAGAATCAAA GTTAGGAAAC ATTGTACCAG AGTACAAAGA AATCAATAAT     1440
CGCGTGAATG TAGCAACAAA CAATCCAGCT TCACAACAAG TTGATAAGCA TTTTGTGCT      1500
AAAGGCCCAG AAGTAAATAG ATTTATTACG CAAAACAAAG TAAACCACCA CTTTATTACT     1560
ACGCAAACCC ACTACAAGAA AGTTATTACT TCATNCAAAA TCAACACATG TNCATTAACA     1620
TGTNNATCAT GCAAAAGGAT TCTTTTAATA NACACTTTAT TGTTACACAT CAGACTCGCC     1680
TAGATATACA CATCCATCTC AATCTTTATT ATCAAGCATC ATGTGCAGCT CCTGGATATC     1740
ACGCGCATAA TTTGGTACAC AGGGCATGCT AGCATTAA      1778

```

(2) INFORMATION FOR SEQ ID NO:95:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1568 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

5 GTGCTTTTGT TCAAGCTTCT GATCAATTTT TTATCGTCTT TTGTAATTTT GCGAATGTCT 60  
TCAAACATTG ATAAGACAAT CTGACCCACA TTTTGTAATT CTTTTTGAGT TTCTTGTAAT 120  
GCAACACCAG GTGCGTGATA AACAAGATCT TTGTGTAAGT GCTGAGGNNT ATAGTCANCA 180  
10 GCNATATCTT TACCTGGGAC AAGCTTTGTA NCTATCCATG CTAAACCTGC TACANATGGT 240  
AATGGAATCA NAGTATTTGT TATGTTGAAG ATACCATGTG ATACTGCAAT CGTCATCGCT 300  
GGGTTTAAAGT GCCATACATC TTGTAACANA CTAATCAAAT GAATCACAAC TGGCAAGAAA 360  
ATTGTGNAGA TAATTACCCC GANTAAATTA AAGATGACGT GTACAAGCGC CGCACGTTTT 420  
15 GCAGCGATTG ANCCGGCTAA ACTAGCTAAG ATAGCTGGAA TTGTGGGACG AGACAATGTT 480  
ATCACCTAGG AACACAGGGA TTGCTGCGTT TNAGCTGATT AAATCTTGTT GATAAAATTC 540  
TTGTAAAATA CCAATCGNCG GACTTGAANT GTTGANCTAG TGCTGTTACC CCTGCGCCGA 600  
CAATGACACC AAGTATTGGA TGTGATGACA ATATCAAGCA TTAATTGNNT AAAATCCATC 660  
20 TAATGATGCT AAGGGGTTTA NCGGCATCCA CCCCATAAAT TCTAAGACCG AAGAAAAGAG 720  
ACCCCGAACC CGATAGTATG CGGCCAATGT TATTGATTTT AAGAGCGTTT AAAAGAAAAA 780  
GATTAAAAAT GCACCTAATG CTAAATTTGG CATTTGCATA TTCGCCTAAA TCTAATNCCG 840  
ATAATAATG CAGTTACCGT TGTTCGGATA TTGGCACCCA TTATCACTCC AATAGCTTGT 900  
25 TTAAACGTCA TAAATCCAGC TGTACCAGT CCGATTGTGA TAACGGTCGT ACCTGAACTA 960  
CTTTGTATTA AAATAGTTAC AACGATACCT GCAATAACAC CTAATACTGG ATTTGATGTA 1020  
AATTGTGTTA AAATATCTCG TAGCCTGTCT CCTGCTGATG CTTGAAGCCC GTCTCCCATG 1080  
ATTTTTAAGC CGTAAAGGAA AATACCTAAA CCACCTAAA AGGAGAAAAT GACTTCTGTA 1140  
30 ACCGACATTT CCATTATTTT CACCTCAAAT AAGCTTTATA TTTAGATTAT CGCTTATAAT 1200  
TGTAATTTA ATGTTAAGAT TAGGTAAAT TATTTAACA TATATGTTAT TTGTANATGA 1260  
CTTGTAATAAT ATCGTCACTT ATTATGTNAA TTTTCAGTGT GAAATGGCAG GTNTGCAATA 1320  
ACGTGTTTAA CAAAATGATG CAATCAATCA TGTAATTATG TTTCATCAA AAAATCATGT 1380  
35 GAGTGGGATA ACGAAATAAA GTTTGTGAAC ATATCATTTT TATCCCCTC CATGATTTGA 1440  
AATCACCAA TAAAAATCTA TTAATGGTTT TCGTTATAAC AATTTGTGTT CTTTAAATAA 1500  
TGTCTCAATG TACGTACCTT TTATCTTTTT AAGGAATCCT GCTAATGCGA GTTCTGTCAT 1560  
TTTCGAAT 1568

## (2) INFORMATION FOR SEQ ID NO:96:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1317 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

55 ACCCCATTGA TTCCTGAATC CCCTGAAAGA AGTTGAATCA CCTTGAAGAT CCTGATTCCN 60

EP 0 841 394 A2

CCTGTACTTT GTGAACCCAC TGAAACTTAT TGAAGAAGAA TCCCCTGANC CTGTCTGATG 120  
 TTGATAATGA TGCCGACACC GATGTGCTTT GTGATGCCGA TGTACTAGCA CTCATTGACA 180  
 5 TTGATGTTGA TATCGATGTA CTTAAGGAAC CAGATGCACT TGTACTTGTT GACTGCCTTG 240  
 TGACATTGAA TCACNTAATG ATGTAGATGT GCTTGTGAG CTCGAGTCAC TTACACTTGT 300  
 TGAACCTGAT ATTGAGTCAC TTAAACTTGT CGATGTTGAA ACTGATACGC TTCCGCTCAT 360  
 TGAGTCAGAT GTTGAAAGTG ATGTACTCGT TGAATTTGAT CCACTGATGT TAGACGAATC 420  
 10 ACTTGTAGAC ATTGAGTCGC TTTCTGATGC ACTGATGCTC ATAGAGTCAA ATTGACTATT 480  
 ACTTGTGAG CTTGACTGCG AATCGCTCAC ACTTGTGAG GTTGATTCTG ATCCACTCAA 540  
 ACTTTGCGAG CTAATCAATG ATTTTGAATC ACTTAATGAA TCCGAAGTGC TAAGACTTGT 600  
 GGAACCACTT AAAGATATTG ATCCACTTAA TGAGTCGGAG TCACTTGTAC TAGTAGAATC 660  
 15 ACTCATTGAT ATTGAATCAC TTAGCGAGGT AGACTCGCTT ACGCTTTCTG AACCCTTAA 720  
 TGATGTTGAG GTACTCAATG AACCAGATGT ACTTGTGAA GTCGAACCAC TTGTTGATTT 780  
 TGAATCACTT AATGAATCAG ATTCCTCAC GCTTCTGAA CTTCTTAGTG ACGTCGATAC 840  
 ACTTAATGAT GACGAATCGC TTGTGCTTAC TGAATCGCTC ATCGATTGTG AGCCACTCAA 900  
 20 TGAACCTGAC TCGCTTACAC TTTCTGATTT TCTTAATGAC GTTGAGACGC TCAATGAGCC 960  
 AGAATCACTG ACACTTGTG AGCCACTCAT CGATTTAGAG TCACTTTCAG AATTAGATTC 1020  
 ACTTCACTT TCTGAATCAT TTACAGATTC TGACATACTT TGTGAATCAG ATATGCTTGC 1080  
 GCTCATTACT TCACTAGCCG ATGTTGATGT ACTTGTGAA TCACTTAACG ATATAGATAC 1140  
 25 ACTCATCGAA CCAGATGTAC TCGCACTTGT TGAGTCTGAT GTTGAATCAC TCACACTATC 1200  
 AGATAATGAC GTTGAATCAC TCATACTTGT TGATGTACTT GTTGAAAGCG ACATACTTTG 1260  
 TGAATCACTA GTACTTGTAC GCATCGAAGT ACTAGTTGAC AGCTGATGTC TCGTGCC 1317

30 (2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2146 base pairs  
 (B) TYPE: nucleic acid  
 35 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

45 CCCAAATTGG CGATAGCTAA TATTGATAA TGATGCGCAT TTCATACAAT TTGTGACTTG 60  
 GCAATTATTG AATATTTATA TAATTTTTTC TCGTAATAAA CACAACCTTG ATAGCGCTAA 120  
 AACAGCTGTG TTAAGTTATG AGTTAACGAA TTTAACACAT TTTACTAGGG CGGCATTTAA 180  
 GAATATTTAT ACGTTTTTAN CGAATATTTA TTTATGTAA NACGCTACCA AAAAGTTAGA 240  
 50 CTTCCCTCCC ACTAAAGTNC CACTTTTTTC TTTCAACTTT TTTAANAAC GGATATGCAA 300  
 CTTTGTAGTAT TGGTATCAAA ATGATTGTTA GGTCATATTC TATCAATATA TTTTCTAAA 360  
 GAATTGCTTT TATTAACCTT CAATTATGTA CCTAACCTAA AAAGAAGCCA AGGCAACGAA 420  
 TGTTACCTTG ACTTCTAATA CATATTCAAC TAACTATATA TTCAATCATA CGCGCATGCG 480  
 55 AGAGTGATTG TTGTACATCT ATAATGCGTT GATTTAAAGA ACCTTTATAT GTTAAATCAG 540  
 GTTGAATAA GTGTTGTATA AATAGACCAT CGACTAAAAC GTCAATGTAT GATAATAACT 600

EP 0 841 394 A2

	CTCGACGTTT	TGTACAATCA	TTTGCTAAAT	ATTTCATATA	AAATCCAGTC	CATACCCAAA	660
	TTGTCTTTGT	ATTTCCAAAA	CGTGCTCGAA	ATGCTTTGAC	AAGATTTAAT	GTAATATCCA	720
5	AATTACAAAA	TGGTTTCGCCA	CCTAATAGAC	TTAGCCCAGA	TATATAATCA	TGATCGCAAT	780
	CATCTAATAT	TTCTGCTAAT	ATTTCATCAG	TGTATTTCTC	GCCATATCTG	AACTTTTGTG	840
	AGGCTTTGTT	ATAACATCCA	ACACAATTAA	ATGGACATCC	TGATACATAA	ACACTGCATC	900
	TTACTCCTTC	ACCGTCAACA	AAGCTATTTG	AATCCTATTT	TAGCAATATA	ACCTTGTCCT	960
10	TGTTTAATGT	CCTAAAAGTG	TCATCCTTTA	GGCGCCTTCA	TATGTTTAC	TCGTGCGCAA	1020
	ATTTCTTTAT	GACGGCCTTT	AATTACTGGA	CGTTGAACTG	GATTGCCTAA	GTAACCCACA	1080
	TGTTTCGTTT	ACGACATCAA	CTGTTTTAGG	ATTATCATTG	NCACAAGTTC	GGGCATTTAA	1140
	ATCCTTTTTT	AGNTGCTTCA	AAATCTCCAT	CCGNNATCAC	ATTCAATAAC	AATGGANCAA	1200
15	ATCGGNATAT	TTGGNNCCTA	AGGTTAACCA	ACTTNGGCAA	TAAGAGTNGG	GCCCCAATACC	1260
	CGNTTCTAGG	GCTTTCAAAT	TGTGTTGCAA	TTTCGGATAC	TCACAATAGT	GAATGAAACC	1320
	ACCACTCGCA	TAATAAGGAT	AATCTTTTTT	AAAATCTAAC	TTTTCAAAAAG	GTGTAACATC	1380
	TTTACGTACA	TCATAATGGA	AAGAGTTTTG	ATAATATCCT	TTATCTGTAA	TGTCTTTAAT	1440
20	ATCTCCAAAT	CTCTCTTGGT	CTAAACGACA	AAAACGATCC	GTTAGCGANN	CACTCGGCGT	1500
	ACTCGTAAAT	ACTGANCCAA	ATGTCATATA	ATTCTGNCCA	TNGCGTGTGA	TAACGTNTCA	1560
	TTTCTTTAAG	AATNACAAGC	GTAAATGCTT	GNGCTTCTNG	AGATGTCTCC	CAGTCTGGAC	1620
	CATAGAAAAC	AGNAGCTGTN	TCATACAACC	CTATCTAGCC	CATTGAAATC	GTTGCACGTT	1680
25	TATTTTAA	TAACTCAGCA	ACATCATCTG	TTTCTTTTAA	TTTATAGTTA	AAAGCGCCAC	1740
	TTTTATATAA	AATTGGTGCG	TTATTCGGTA	CAGNATCTTT	CAAACGATTT	ATACGATAAA	1800
	GTAATGCATC	ATGTAACACA	TCGATACGTT	CATAAAAGAT	TTCCCAGAAT	TCGTCATAT	1860
	TACCGGCAGA	TTCTAATGCC	ATTCTAGGTA	AATTAAGTGT	NACAACACCA	AGATTACAAC	1920
30	GACCATTATT	TTCAAAATGA	CCTTCCGCAT	CTTTCCAAC	TGGGTAAAAA	TGAACGACAA	1980
	CCCATTGGNG	CTTTGAAATC	ACCTAATATT	TCTACGAGTT	TGGTCAATAA	TTTAAATAT	2040
	CTGGATACAT	ACGTTTCCGN	GGAACAAC	TAATGCTAGT	TGTTTAATGT	CATAGGTCGG	2100
	ATCTTGCGGG	ACTAAAGTTG	GGGTCCNTC	TCTATTGGAA	ATCAAA		2146

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1015 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

50	ATTGTCCCAC	ACCATTACT	ACAAGANACG	AGCGTCCTCC	ACATTGTTAC	NTTGTGCNAT	60
	CAATNCTTCT	NGTTTGTTTT	GAANNNAAGA	CCGAAGCATT	TGTANTCTGC	TGTGTTTCGCA	120
	GCCTTGTTTC	GCACGTCTTA	AGGTATTGAA	TACCGTTTAA	TGCAGTATTG	GCTTGTGTAA	180
55	CAGCTTGCAT	TGCTCTATCG	ACATCTGCNT	TAGGTGTATT	TCCTCCAGCT	GTTTTATNTA	240
	AAATAGTTGC	TGCTGCATTT	ACTGCTTGAG	AATAAGCCGT	TCGTTTAGCA	TCATCAGCAT	300

EP 0 841 394 A2

CTTGATAATT TTGACTTTGT AACGTCGTGT CTTTATCAGC AATTGATGTT TCTAATTGAC 360  
 CCATAGCACC ATCTAATTGT TCGCCTTTGG CTTTAACTGT ATTAACACCT CAACATTTGT 420  
 5 TGCTNGTGTA ATTTCAATTAT CTAACGCATT ACGTTGTGCA TTATTAATGT GTGTTAATGT 480  
 ACCTAACGTT TGTTCGCAG CAGCTTTAGC TTCATTTAAT TTCGCATCAC CGCNCAACGC 540  
 CGTCTTCGTA CTGTNCACAT TGTGTAATGC TTGNNCAACT GCTGCTTTGT CTACATTGTG 600  
 ACCACTAGCT TTTGTTAAAA TTGCTTTTGC TGCATTTTACT GCTTGATCAT AAGCTGATTT 660  
 10 CTTACTTGGC TCAGCATCTA GGTATTTCTG AGTTTGTTTT GGTGTGTCT CATCATTTGAT 720  
 ACCATTTTGT AAACCTGTGC ATTGCGTTAT TTAATTCTTG TTGCTTTNGC AGGTTCTTGG 780  
 ATTTACACCA GCTACTGTAG GTGGNACCAT CAATGTTAAG CGTTAACGCA TCTTTTTGTG 840  
 CATTGTTAAT TGATGTTAAG TTATTCAAGT TTGTTTTTCG CTGTTGTCTT AGCTTGAGCT 900  
 15 AAGTTTGGC GCACCATTTA ATGCATGTTT TTTAGTNGGT CACTTGTGAT GTTGCTTGGC 960  
 GTAATAGTAT TNGGGNTCCA TCCNNTGGNT TCCACGTTG GATTAATGAT TGCTT 1015

(2) INFORMATION FOR SEQ ID NO:99:

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1550 base pairs

(B) TYPE: nucleic acid

25

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

GGCACCAGCT AGAGCGAGTT TTTATATATA CAACACGAAA GAAGACATTG ATCAGTTAAT 60  
 35 AAATGCCTTG AAACAAACGA AGGAGTTTTT CTCTTATGAA TTTTAATAAT CTAGATCAAT 120  
 TATATAGATC TGTCATTATG GATCATTATA AAAATCCTAC AAATAAAGGT GTATTANATA 180  
 ACGGTCTAT GACAGTAGAT ATGAATAACC CGACATGCGG TGACCGTATA CGACTAACAT 240  
 TTGATATAGA AGACGGCATT ATAAAAGATG CTAAGTTTGA AGGTGAAGGT TGTTCGATTT 300  
 40 CAATGGCAAG TGCATCGATG ATGACACAAG CTGTTAAAGG GAATCCANTT GGAGAAGCAA 360  
 TCCAAATNGA CCAAGGAAT TTACCGNAAA TGATGCTTGG TTGGAAGCCT ANTGTGATNC 420  
 AGGGAAGAAT GGGGAGATAT TGAAGCATTC CNGGGTTGAT CTCAATCCCA GCTCGTATTA 480  
 AATTTGTCCA CATACTTGG AAACATTGGA AAAAGGGTCC TTGTTCTTAA AGGAGGGTAA 540  
 45 AACAGGAGGG TCCGGCTTGA AGAAGAAAAG ATCCTGTAA TCATAAGATG ATTTTGATAT 600  
 TAAGACATAT NNAAGTATNN NAATTTTTAA TAAAGATGTC ATGTCATTGT AATAAATATG 660  
 GTTTACATCA TTGAATTAAA AACTTACGCA CCGCCGTGT AAATATATTT TTAAGGAGTG 720  
 ATTGAAATGG CTAATAAAGC ACCTGATGTT GGGGATTATA AATATGGATT CCCC GCCGAT 780  
 50 GATGTATCCA TTTTCAGATC AGAACGTGGT TTAACGAGA ATATCGTTAG AGAAATTTCT 840  
 AACATGAAAA ATGAGCCGGA ATGGATGTTA GATTTCCGTC TTAAATCATT AAAATGTTT 900  
 TATAAATGC CAATGCCCTCA ATGGGGTGGC GACTTATCAG AATTGAATTT CGATGACATT 960  
 ACTTACTATG TAAAGCCTTC AGAACAAGCT GAACGTTTAT GGGATGAAGT GCCAGAAGAA 1020  
 55 ATTAAAGAA CTTTCGATAA ATTAGGAATT CCTGAAGCTG AACAAAAATA TTTAGCTGGT 1080  
 GTTCTGCTC AATATGAATC TGAAGTTGTT TACCATAATA TGAATAAAGA ACTGAAGAA 1140

EP 0 841 394 A2

AAAGGTATTA TCTTTAAAGA TACAGATAGT GCTTTACAAG AAAATGAAGA ATTATTCAAA 1200  
 AAATACTTTG CTTCTGTAGT ACCTGCAGCA GATAACAAAT TTGCGGCGTT AACTCAGCA 1260  
 5 GTATGGTCAG GTGGNTCGCT CATTATATGTA CCTAAAAATA TCAAACCTAGA TACGCCACTA 1320  
 CAAGCTTATT TCCGTATTAA CTCTGAGAAC ATGGGTCAAT TTGAACGTAC ATTAATCATT 1380  
 GCTGATGAAG GTGCTTCTGT ACATTACGTA GAAGGTGGT ACTGCACCAG TTTATACAAC 1440  
 TAGNTCTTTA CACTCTGCTG TTGGTGGNAA TCATTGGGCA TAAAGATGCG CACNGTCCGC 1500  
 10 NTTATTCTTA CGAANCAAAA CTGCGGGACA ATGTTTTCAA CTNNAGGTAC 1550

(2) INFORMATION FOR SEQ ID NO:100:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1449 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

TGATAATACT GCTTAATACA ACATTTATTG CAATTAAAAG TGCAGTAATA GCCAGTTTTT 60  
 CNTTGANTTC AAAATGANTG TCCTCNCNTT TTGTTTGTA NTAACTACTA TGCTTGGCTT 120  
 30 TATTATGGTC ANNTAAACGT GTTTCATTG GTTGATACTA ACATTTTCAA TAATTGATTC 180  
 GCTTCATATT GTGAAGTTTG AACTGTTC NCTATGGGCA ATGTATTTAT TTCTGCTTCT 240  
 ATACTTGAN TGGTATGTTT CGACTGTTCT AGCGCATTTT GTTCCCGTA ATTTTGAAAG 300  
 TTTACTGCTT GTTTTGTAG CTTTTTAAAC ATATCCATT TAGTCTTTAT CGTTTGGTTC 360  
 35 TGATGAATCT GTGCTTCAAT TTGTTGATAT GTTTTGTAG AANCCCAAAT TTTAATTTTA 420  
 NTTGCAATAT AACCCGCTTG TTCCAAANCG TCATCTTTAT AAAACAATTT ATTTGAATCA 480  
 CCNATTTCCG GCTCCTACTT CCTTTACCAA ACTTGCCATT TAATGGAATA CTGTTTAGCT 540  
 TCATCTATTC GACTTCCAC TAGTTTACCC AATCATTTCT TTAGGTGCTT TGAAATTAAC 600  
 40 TAGCTTATTT TTATCAGTGT AGCCAGCAAG AACCTGATCA TCTTTTTTAC TACTACCTTC 660  
 ACAAAGTACT GTTACAGTTT GTCTTCGTA CTTACTCATA GCTATTTGTG AATAATGACC 720  
 AACTTTTTTA TTCAAACGTT GCAATCGTTC CTTTTTGACA TTTAAAGGTA CATTATCTTT 780  
 CATTTTAGCA GCAGGCGTAC CATCACGTTG TGAATACAAG TACGTATATG CATGTTCAAA 840  
 45 ACCAACTTCA TCATACAGAG TTAAAGTTTC TTCAAATTGT TCCTCTGATT CATTGGGATA 900  
 CCCTACAATA ATATCTGTAG TTAATGCTAC ATTAGGAAGT CTATCTTTGA TTCGTTTTAC 960  
 TAAATCCAAA TAACTTTCTC GTGTATATTT TCTACCCATT ATTTTAAATA CTGCATTATT 1020  
 TCCAGATTGA ACTGGCAAGT GGATATGAGG AACGATATTA CCACCCTCTG AAATAACATC 1080  
 50 AATCATGTGA TCTGTAAAGT CCAAGGATG ACTTGTTGTG AAACGAATC TTGGAATCGC 1140  
 TATTTTAGAA ATTGCTTGTA AAAGATCTCC TAAGTCATAT TCTATATCCT GTAAATCTTT 1200  
 ACCATAAGAA TTTACATGTG GACCTAAAAG CGTTATTTCT TTGTAACCTT CACGAGCAAG 1260  
 TTCACGTA CTATCTATAA TGTCTTCAGG TCTACGGGTC CGGTCTTAC TCCTTGTTNAA 1320  
 55 TGGAACAATA CAATATGTAC AAAACTTATC ACAACCATAC ATAATATTGA CCCATGCTTT 1380  
 ATGTGCTTTC ACGGACTTGT GGNAGATTAC AATAACGTCT NCTTCTTTTA GACATACTCA 1440



CAACAATGT

1449

## 5 (2) INFORMATION FOR SEQ ID NO:101:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 459 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## 15 (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

20 GAAAGTGTTC CAGAACGTGG AGCAAAAGGC TTCGGAAGTA GCGGAGTGTA AAGACATCTT 60  
 AGATCGAGTT AAGGAGGTTT TGGGGAAGTG ACGCAATACT TAGTCACAAC ATTCAAAGAT 120  
 TCAACAGGAC GACCACATGA ACATATTACT GTGGCTAGAG ATAATCAGAC GTTTACAGTT 180  
 ATTGAGGCAG AGAGTAAAGA AGAAGCTGAG CGCAAATACG AGGCACAAGT TAAGATAAGG 240  
 25 AGAGATGGAG ATGCCAAAGA AAACGGNAAC GATTGATGTA GATGAAAAC TATTAGTAGT 300  
 AGCTAGTAAT GAAATATCAG AACTATTATA TGAATATGAC AGTGAGTTAA TCNCAGCTGG 360  
 ATGAAGATGG CGATAATAGA GATATCGGAA GGAAAAAAGA GGACGCATTA AAACAAGCTA 420  
 TACAAAATTT CTCGATAAAT TACATGGGGG TGTTTAGTG 459

## 30 (2) INFORMATION FOR SEQ ID NO:102:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2005 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## 40 (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

45 GGCACGAGCT CATAATCGC TAGTAAGAAT AATAATCTTA GTATTTGTTT AGTCATTCCC 60  
 CACACTCCCT TATATTTTCA AACAACTGAC TCACTTTAGC ATAATTCCAT CTCTCCTATC 120  
 TTAACCTTGT GCCTCGTANT TGCCTCAGC TTCTTCTTTA CTCTCTGCCT CAACAACTGT 180  
 50 AAACGTCTGA TTATCTCTAG CAGTAGTAAA ATGTTTCATGT GGTGTGCTT GTTGAATCTT 240  
 TGAATGTTGT GACTAAGTAT TGTGTCATTC CTCATAGCTC CCTTGAACCTT GTTTGAGCTT 300  
 ACTCATAAAA AACATTACTA AAAATGCTAT TAAGATATGC GTCTTTTGAT GTTTATAAGC 360  
 AAATGTAGAT ATCATAAAGA TAGTAGCAAG CATTAACATT TCATATATGT TTGTGTGTAT 420  
 55 AGTCTTTTTA CTCTTAAGAA AAATAATTGC TATGCGATAA AAGAGATAAA CGCCAAACCC 480  
 TATTAATAAT ATTTCTAACA TGTCGCTCAC TTCCCCAAA CCTCCTTGAC TCGATCTAAG 540

EP 0 841 394 A2

	ATGTCCTTTAT	ACTCCGCTAC	TTCCGAAGCC	TTTTGCTCCA	CGTTCTGAAA	CACACTCGAA	600
	TTCTCTCCACT	TGCTTTAGTT	CAGGTGTCCA	TATAGGCACG	ATAACCAATT	GAGCTAGTTT	660
5	GTCGCCTTCG	TTGATTGAT	AAGTTCATA	TTGTCTTATG	GCGTCACTCA	AATCGATTTC	720
	TCCTTTAATA	TCAAAAACAC	CTGGTGTGAT	ATAACCATTC	GATGCAATAG	CGTCATTCTT	780
	GATATTAATC	CCTAAATTGC	CGTGATATCC	CGCGTCTATC	TTGCCTGTTT	CAATCACTAA	840
	ATGCGTTTTA	CTACTTACAC	CACTACGACT	AGTTAATAGT	CCGACATAGC	CCTCTGGTAT	900
10	ACTCACAGCT	ACATCTGTTT	TGATCACTGC	TTTTTCTTGT	GGTTCGAGTA	CGACAGTTTC	960
	AGCTGAGAAT	ATGTCATAAC	CTGCATCCGT	CTTATGATTT	CGTTCCGGCA	TTCTAGCATT	1020
	TTTTGATAAT	AGTTTTACTT	GTAATGTGTT	AGTCATTTTC	CTATTCCTCC	TCATATTTAT	1080
	AGACAACCTG	ACCTGCCATA	ATCCCTACTG	CTTCATCAAG	TTCAATACCT	CNTTTTAACTG	1140
15	AATGTTGAAT	AGCATTTGTC	ATTCCCTCAA	GTATTTTCATC	AAACGCTTGC	GCTTTCTTAT	1200
	ACACGTCCTC	AATCTCTTTT	AGCAACCCCT	CTGTGTCATT	ACCGTATACG	CACTAGCACT	1260
	AATAACGGAG	TGTTCCGATT	TTTCGCGATT	ATTCAATTGGT	GTCATCCTCC	ATAAAAAATT	1320
	TATTGTTTAA	TTCCATCCCG	AATTTAACTC	TTTCATCATC	GTTACCGAAT	TCGTTTATTA	1380
20	AATCTTTTTT	AACGCTCTTG	CAATACCTAT	CCCATGCGCT	TGCTTTCTTC	TCCAGTTCTT	1440
	TGTTACAATC	TCGTAACTTC	GCTATATCCC	CAATAAGCTC	ATCTCGTTGC	TTCTTGTACT	1500
	CTTCACGATC	TTTTAATGCT	TTGTGAAGTT	TATCTAATAA	CTTGTAGAG	TTAGTACAAA	1560
	GATTTTTATA	TTGTTTCATC	GATAAGGTGA	ACGTCATCTC	ATAACCTCCA	ATAGCATCTC	1620
25	ATTTTCAAAA	ATATTTCCAA	CAATTTCAAT	AATATCGGCA	TTTTCACTTA	GTAATTCAGT	1680
	TACATGCTA	AAAGTTATAT	AAAAGGCTCC	TTCTTTAAAC	TCGATAAAAC	TTACTTCTCT	1740
	CGAATTAACA	ATCTTGGAAC	AATATCCCCT	TCATAAATCT	CCACAACCGG	GCACATCTTT	1800
	TAAATCCTGT	GTATTGGTAA	TAGGTTTTAC	TTCAATTGAA	ACTTTTATTA	ACCTGGTGGA	1860
30	AATCAAAATG	TACCCACTTA	TTAAAAATCG	GATTTGCGNC	AATAATACTC	AATAACTTNN	1920
	NNTATCTTNA	TCCCAAGCTT	TTAATTTCAA	CATCAATCTT	ACCAACTCCC	CATCTTTCCA	1980
	AATCAATGTC	AACCGGCAAN	GTACAC				2005

35 (2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:

	(A) LENGTH: 796 base pairs
40	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: single
	(D) TOPOLOGY: linear

45 (ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

50	GGCACGAGCT	GTTTTTTTATT	TGCTTCGTTA	CAAAGCATTA	TTGAATTTAT	TTTACGTGTT	60
	CATATTTTGA	AACATCAAAG	CCGTCTTGCT	TAGCTTTGTT	GATAATGTCT	TTGATTGAAT	120
	GTAGTCCTTT	ATCGGCAGAA	TATGATCTTA	AGTTGTCTTT	TGTAGCTTGG	TCAGCATTCT	180
	TATCTAATAA	CACATCGATA	TAGCTTAATT	CATGTCTTAA	GAAGTTTGCG	TCATCATGTA	240
55	GTACGAGTCC	ATTTTGAGAA	TAACTTTTCG	CATCTGCTTG	ATTACCATAT	CCAACAACGC	300
	CAGTTGCTAA	TACACCTACC	ATTGCCGTAG	CTACTAAAAC	CTTTTTAAAT	TTCATATCTA	360

EP 0 841 394 A2

5 TCACCTCCTCT AAAAAATTGTA CTCTATCATA ACACTTGAAT ATTAAGAAAA TTACGGNTTA 420  
 TTAAGTCGGA CTTNANTAAT TCTTAATAAA TAGNTAACT GACAAATATT NGCTNAAATG 480  
 CAATTANTCT TNAAAACCGG GGTATTATGGA TTTTTCCTAC TAAAACCTTG ATTTCAAAAA 540  
 GGGTTTANCT CAAATGAAAC AATAATAAAA AATAATGCAA CATAATAATA AGTACAAATT 600  
 TANTTAAGAA ATTAAATTGA TTGTATATGT ATATTNTGGT AACGTNAAAG AGAAATATNC 660  
 AANATAATTA ATTATTTATT TGAAAAGAGA ATATTAATGA AGTATTAAAC AAAGAGACGT 720  
 10 GAAACGATGC GATATTTAAA AAGACTTTCA TGGNACATAA GCATCTTAAT TTTAATAGGT 780  
 GGTATTGCTG GGTGGG 796

(2) INFORMATION FOR SEQ ID NO:104:

15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1122 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: Genomic DNA

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

AACTTTGAGT NTTTTGTTCN ATATGGAAC TTAAGATTT GGAAATTATN TAGATGGAGA 60  
 30 TATATCATAT AATCCAGAGG TGCCNAGTTA TTCAGCTAAA TATCAATTAN CTAATGATGA 120  
 TNACAATGTA AAGCAATTAC GTAAAAGATA TGACATNCCA GCAGAATAAA GCACCCAAAG 180  
 TTATTTGTTG AAAGGGTACA GGGTAATTTN AAAGCCTCAT CAGTTGGATA TAAAGACAGT 240  
 GAATTTACTT TCGTAGAGAN NAAAGGTGAA AATATATACT TTAGTGATAG TCTACATCTT 300  
 35 GANCCAAGTG AGGATAAATA AACGTGACCA ATAAAGAGTA TGAAATCGAA CCCGGAAAAA 360  
 GAGAGTGTGA AATGATGAAA CGATTAAATA AATTAGTGTT AGGCATTAGT TTTCTGTTTT 420  
 TAGTCATTAG TACTACTGCT GGTGTGGCA TAGGTAAAGA AGCGGAAGTT AAGAAAAGCT 480  
 TTGAAAAAAC ATTGAGTATG TACCCTATTA AAAATCTAGA GGATTTATAC GATAAGGAAG 540  
 40 GCTATCGTGA TGATCAGTTT GATAAAAATG ATAAAGGTAC ATGGATTATA AATTCTGAAA 600  
 TGGTTATTCA ACCTAATAAT GAAGATATGG TAGCTAAAGG CATGGTTCTA TATATCGAAT 660  
 AGAAATACCA AAACAACAAA TGGTTACTAC TATGTCGATG TGACTAAGGA CGAGGATGAA 720  
 GGAAACCGC ACGACAATGA AAAAAGATAT CCGGTTAAAA TGGTCCGATA ATAAAATCAT 780  
 45 TCCAACAAAA GAAATTANAG ATAANAACAT AAAAAAGAA NTCGAAAAC TTAAGTTCTT 840  
 TGTTCATAT GGAACCTTTA AAGATTTGTC GAAGTACAAA GATGGAGATA TTTCATACAA 900  
 TCCAGAGGTG CCAAGTTATT CACCAAAATA TCAAGTAACT AATGATGACT ATAATGTAAN 960  
 ACAATTNCGN AAAAGATATA GATATACCGA CCGAATAAAG CACCANAGCT ATTGTTGANA 1020  
 50 GGTACAGGGA ATTTAAAAGG TTCATCAATT GGCTACACAA AAATTGAATT TACTTTCGTA 1080  
 GAGAAAAAGG GAGAAAATAT ATACTTTAGT GATGGGCTAC AG 1122

(2) INFORMATION FOR SEQ ID NO:105:

55

(i) SEQUENCE CHARACTERISTICS:

# EP 0 841 394 A2

(A) LENGTH: 652 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

CGGTGGTCTT	CTATTCCCGT	TTTGCCATCT	CNAATTTCTT	ACCATGGCCC	ACAAAGTATT	60
GTNACGTTTG	AGCTGATCCT	TATTTAATTT	TNAAATGAG	AAAAATGTCT	CCAAATCCTG	120
TACCTNCCCC	ATATTCCCGC	AATACAAATA	AAATTTTGTC	GTATTTGCTC	NCCGGTAAGC	180
TTTAAATGTG	GTCATTATAA	GANCGGATTG	TCTTGTAATT	GACCGCATGT	CATACCAGTT	240
AGGAATCACA	TGGATATTGT	CAGCATTTTT	AAGAAATTTG	ATGATTTAGT	AAGTAGTTTT	300
TCATTTCCGT	ACCCAAGGAC	AATGACATTT	TCAGCATTCT	TGTAGACATG	TCTATTAATG	360
TAACGCATCA	GCTTATCAAT	CATGCTACCT	GGACGAGTTG	CACCTGTCCT	AATCGCATTG	420
TCAGGTGCTA	TATCATACAC	CACAAAAGAA	TATTTTCTCT	TAAGCAGTCT	GTGTAAAACG	480
TCTGGTATTA	ATGGCAAGAT	TGGTGGATTA	GAGTAAACAA	GAATCTGATC	ATATTTCAAC	540
ATTTTAGGTA	TATTAATCAC	GAATTTTGAA	AATAAACTAA	AGAAATTGAT	GATCCTTCCA	600
ACCTTACTTT	TGTTATTAAA	CCTCGAATAC	TTGAGACGTC	GAATGCGAAN	TC	652

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 392 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

GAATTCCTT	CAAAAGCACC	TTTGATACCG	AATAAATTAT	GGTTTGGTGA	CTTAGATAAA	60
GAAGTATGCGG	GAAGATTTTA	ATCGCAAGAC	TGGTGCGAAT	GTGAAGTATA	TTGAAGCACC	120
TTATGAACCG	CATAAGTTTG	TGAAAATGGT	GAAGGATAAA	GAATTAGCTG	ATGAAAAAGA	180
AGGCGGCTTA	CGTNGTACCG	CTTGTTTTGA	AATGCGTTTG	GATATTGTAG	CGAAAGCAGC	240
TGTAGAACAT	GGCTATGATT	ATTTTGGCAG	TGCAATCACG	TTATCACCTA	AAAAGAACGC	300
ACAATTAATC	AATGAACTTG	GTATGGATTG	TCCAAAAAAT	ATACGATGTG	AACTTATTTG	360
CAAGTGATTT	TAAGAAAACT	AAGGTATTGA	GC			392

(2) INFORMATION FOR SEQ ID NO:107:

(i) SEQUENCE CHARACTERISTICS:

# EP 0 841 394 A2

(A) LENGTH: 3797 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

	TCAAAATGCAG TCAGGGAAGC AATAGGACGA TATGCATAAA GGAGATGGTA AAGTGGAACA	60
15	GTGACAGAAG GTAAAGACAC GCTTCAATCA TCGGAGNCAT CAATCAANCA CAAAATAGTA	120
	AAACAATCAG GAACGCAAAA TGATAATCAA GTAAAGCAAG ATTCTGGAAC GACAAGGTTTC	180
	TAAACAGTCA CACCAAAATA ATGCGACTAA TAATACTGAA CGTCAAAATG ATCAGGTTCA	240
	AAATACCCAT CATGCTGAAC GTAATGGATC ACAATCGACA ACGTCACAAT CGAATGATGT	300
20	TGATAAATCA CAACCATCCA TTCCGGCACA AAAGGTATTA CCCAATCATG ATAAAGCAGC	360
	ACCAACTTCA ACTACACCCC CGTCTAATGA TAAACTGCA CCTAAATCAA CAAAAGCACA	420
	AGATGCAACC ACGGACAAAC ATCCAAATCA ACAAGATACA CATCAACCCG CGTGCCTCAA	480
	ATCATAGATG CAAAGCAAGA TGATACTGTT CGCCAAAGTG AACAGAAAAC ACAAGTTGGC	540
25	GATTTAAGTA AACATATCGA TGGTCAAAAT TCCCAGAGA AACCGACAGA TAAAAATACT	600
	GATAATAAAC AACTAATCAA AGATGCGCTT CAAGCGCTA AAACACGTTT GACTACAAAT	660
	GCAGCAGCAG ATGCTAAAAA GGTTCGACCA CTTAAAGCGA ATCAAGTACA ACCACTTAAC	720
	AAATATCCAG TTGTTTTTGT ACATGGATTT TTAGGATTAG TAGGCGATAA TGCACCTGCT	780
30	TTATATCCAA ATTATTGGGG TGGAAATAAA TTTAAAGTTA TCGAGGGAAT TGAGAAAGCA	840
	AGGCTATAAT GTACATCAAG CAAGTGTAAG TGCATTTGGT AGTAACTATG ATCGCGCTGT	900
	AGAACTTTAT TATTACATTA AAGGTGGTCA CGAGCGTAGA TTATGGCGCA GCACATGCAG	960
	CTAAATACGG ACATGAGCGC TATGGTAAGA CTTATAAAGG AATCATGCCT AATTGGGAAC	1020
35	CTGGTAAAAA GGTACATCTT GTAGGGCATA GTATGGGTGG TCAACAATTT CGTTTAAATGG	1080
	AAGAGTTTTT AAGAAATGGT AACAAAGAAG AAATTGCCTA TCATAAAGCG CATGGTGGAG	1140
	AAATATCACC ATTATTCAC TGTGGTCATA ACAATATGGT TGCATCAATC ACAACATTAG	1200
	CAACACCACA TAATGGTTCA CAAGCAGCTG ATAAGTTTGG AAATACAGAA GCTGTTAGAA	1260
40	AAATCATGTT CGCTTTAAAT CGATTTATGG GTAACAAGTA TTCCGAATAT CGATTTAGGA	1320
	TTAACGCAAT GGGGCTTTAA ACAATTACCA AATGAGAGTT ACATTGACTA TATTAAAACG	1380
	CGTTAGTAAA AGCAAAATTT GGACATCAGA CGATAATGCT GCCTATGATT TAACGTTAGA	1440
	TGGCTCTGCA AAATTGAACA ACATGACAAG TATGAATCCT AATATTACGT ATACGACTTA	1500
45	TACAGGTGTG TCTTACATA CTGGTCCATT AGGGCACGAA AATCCTGCCG AATTAGGCAC	1560
	GAGACATTTT TCTTAATGGA TACAACGAGT AGAATTATTG GTCATGATGC AAGAGAAGAA	1620
	TGGCGTAAAA ATGATGGTGT CGTACCAGTG ATTTTCGTCGT TACATCCATC CAATCAACCA	1680
	TTTATTAAATG TTACGAATGA TGAACCTGCC ACACGCAGAG GTATCTGGCA AGTTAAACCA	1740
50	ATCATACAAG GATGGGATCA TGTCGATTTT ATCGGTGTGG ACTTCCTGGA TTTCAACACC	1800
	GTAAGGTGCA GAACCTGCCA ACTTCTATAC AGGTATAATA AATGACTTGT TCGGTGTGGA	1860
	AGCGNCTGAA AGTAAAGGAA CACAATTGAA AGCAAGTTAA ATTCATCTTC TGAATTTAAT	1920
	AGGCTATGTA AATCGTGCTG TTATCATGGC ACATCAGATA TAAGTAGCAT CACAGTGTG	1980
55	AATCTCAAAA TAGTAAAGTG AAATAAAGCG CCTGTCTCAT TAGCGAAAAC TAAAGGGACA	2040
	GGCGTATCTG TTTATGAGCT TAATAAATTG TATGAATAAT ATGGTTGATC GAATAACTGT	2100

EP 0 841 394 A2

```

5  TTATCATTGA TGATAAAATTT GAGTTTTTTTA AAAATAATTG ATATATTACA CCATTGTTAT 2160
   AGCGTTTAAA GAAATCAACC CAACTTTACG ATAAATAGTG ATTGCTTCGT CATTAGGTCT 2220
   ACGATCAAAA TCATGCTCGT TTTTATTCAC GCGTTCAAAT GTTGAATGTG GAACATGATT 2280
   CATGATATGT TCGCTTTCCCT CAACGGGAAC ATCATAATCG CCATTACAAT GCGCAATGAA 2340
   AACAGGTGGA AGTGTTTTAA GNTCATCTGG TGCAATATTA TATTTTGAAT CAGTATAATC 2400
   ANCAATGTTA ATCATATTTA TCCATTTACC TGTGCCACGT GCATAAACGT AGAGTAAAAA 2460
10  ACGTGTGCGA TTTGATCTTG ANCAACCGGT GTTGGTGAAG TGAGTTGTCC AATCATTGTT 2520
   TCGTTTATGC TTTGAGCTAT TTTTGCGTAA TACCTATTAG TTGTTTAAA AGGGTTCAGT 2580
   GTTGATGCGA CTATAACCAT AAAAATCAAT AACACCATCA ATATCTCTGT CTCGTGCAAT 2640
   TAATAAGACT TAAATATGCA CCTGATGATC TGCCAAAGGT AAAAATAGGG CAATTAGAAT 2700
15  ATTGTGATTG AATCGCATCG AATGATGCGT AGACATCCTC AATAATGCAA TCGAGACTTA 2760
   CTTCTGGTAA TAAACGATAA CTTAGTTGAA TTAAATCGTA ATGTTCCGTA AGGATATCGA 2820
   TATACTGTGG GGATAAATCG TTAGCTTTAC CGAACATTAA TCCACCACCG TGGATGTAGA 2880
   CAATAACGCC TTTTGTGGT TGATTTTTTG CTTTAATAAT TGTGTAAGGT AATGCAAATG 2940
20  CATCTTTAGT AATTACTTTA TATTTAATTT CAGTCACGAT TTAATAGGCT CCTTAGGAAT 3000
   CCGATATGTA TGTCAATTATA ACACGTGTCNT NAATTTCCAT GNAAAATAGT CTTAAGACGA 3060
   TGAGTCATGA TAATTCGTGT CCAATGACG TAAAGCGTCN CGGGTATGCT TCTTTAGACC 3120
   TTCCCCATAA TCCATCATTT TAACAATATC TTTAAAAGCA GCATGTGGNA TGGCTAAATC 3180
25  TTCTAAATCT GCCATAGAAA ATTCAAGATT GATATCATGT GGTGCTGTT CAGCAAGTTT 3240
   ATGCACAAAG TCAGGTTCTG TGACCAAAGG CGAAGACATG CCGACCATAT CTGCATGTTG 3300
   TAAAGCATCT AAAGCAGACT CTGGAGAATT AATCCCGCCA CTTGCAATTA AAGGGATACG 3360
   ACCTGCTAAA TGTTCATAGA CAATTTGGTT AACTGGTCGA CCGAAATGAT CACCTGGTGT 3420
30  ACGAGACGTA TTTTGATAAA TATGTCGACC CCAGCTAGCG ATTGCTAAGT ATTGGATGTT 3480
   TGAAACGTCC ATGACCCAAT CGATTAAATTG GTTGAACCTG TCAATGGTAT ATCCTAAATC 3540
   ACTGCCTCTG GTTCTTCTG GCGTTGCTCG AAATCCTAAA ATAAAATTGT CAGGTGCTTC 3600
   TTTATCAATC ACTTCTTGTA CCGCACGCAT AACTTCTAAA CATAATCTTG CACGATTTTT 3660
35  TAATGAGTCG GCACCGTAAT GGTCTGTACG TCTATTTGAA AAAGTTGAGA AAAATGTTTG 3720
   AATCAGCAAA CGTTGTGCAA TCGAAATTTT CACACCATCA AAACCTGCTT TAATCGCGCG 3780
   TGCATCGAGC TCGTGCC 3797

```

40 (2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:

```

45  (A) LENGTH: 3165 base pairs
     (B) TYPE: nucleic acid
     (C) STRANDEDNESS: single
     (D) TOPOLOGY: linear

```

50 (ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

```

55  CAAGCCAGTT CAACTNACCG ACCCGGTATA TTTGGGGAAT ATACGAANCN CNNTCAGATG 60
     ATCNAAATAT GANTGANACN CGGTNGCGTT CTCNAGNCAN AACAGNGTTA AATAANCGTA 120

```

EP 0 841 394 A2

	AGTCGGCAGT	TATGTCANAT	GAGATTACTT	CTNCTGATTG	GGNAGTTACG	NNTCCACACG	180
	AGANNATATC	ANNTGGTGAN	NCAGNCAGAG	TAAAACATAG	AGATTTTGAC	CCGCCATTGT	240
5	ATGTAGAGGC	AGAAGTTATT	GCCGAAGAAT	ATAACATAAT	TTCAGAAAAT	AGCACATATA	300
	CATTCCGGTCA	ACCTAAAGAG	TTCAAAGAAT	CAGAATTACG	AGAAGAGTTT	AACAAGCGAT	360
	TGAACATAAT	ACATCAAAAG	TTAAACGATA	ATATTAGCAA	TATCAACACT	ATAGTAAAAG	420
	ATGTTGTAGA	TAGTGAATTA	GAATACTTTG	AACGCCAAAAT	ACACAAAAGT	GATACACCGC	480
10	CAGAAAATCC	AGTCAATGAT	ATGCTTTGGT	ATGATACAAG	TAACCCTGAT	GTTGCTGTCT	540
	TGCGTAGATA	TTGGAATGGT	CGATGGATTG	AAGAAACACC	AAATGATGTT	GNAAAATTAG	600
	GTGGTATAAC	AAGAGAGGAA	AGCGCTATTC	AGTGAATTAA	ACAATATATT	TATTAATTTA	660
	TCTATACAAC	ACGCTAGTCT	TTTGTCAAGA	AGCTACAGAA	TTACTGGAAT	AGCGAGTACT	720
15	TAGTAGATAA	TGATTTGAAA	GCGGACTTAC	AAGCAAGTTT	AGACGCTGTG	ATTGATGTTT	780
	ATAATCAAAAT	TAAAAATAAT	TTAGAATCTA	TGACACCCGA	AACTGCAACG	ATTGGTCGGT	840
	TGGTAGATAC	AAAAACTTTA	TTTCTTGAGT	ATAGAAAGAA	ATTACAAGAT	GTTTATACAG	900
	ATGTAGAAGA	TGTCAAAATC	GCCATTTTCA	GATAGTTTAA	ATTATTACAG	TCACAATACA	960
20	CTGATGAAAA	ATATAAAGAA	GCGTTGGAAA	TAATAGCAAC	AAAAATTTGGT	TTAACGGTGA	1020
	ATGAAGATTT	GCAGTTAGTC	GGAGAACCTA	ATGTTGTTAA	ATCAGCTATT	GAAGCAGCTA	1080
	GAGAATCCAC	AAAAGAACAA	TTACGTGACT	ATGTNAAAAC	ATCGGACTAT	NAAACAGACA	1140
	ANGACGGGAT	TGTTGAACGT	NTAGATACTG	CTGAAGCTGA	GAGAACGACT	TTNNAAGGGG	1200
25	AAATCAAAGA	TAAAGNTACG	GTTANACGAA	TATCGAAACG	GATTGGAAGA	ACAAAAACAA	1260
	TATACTGATG	ACCAGTTAAG	TGATTTGTCC	AATAATCCTG	AGATTAAAGC	AAGTATTGAA	1320
	CAAGCAAATC	AAGAAGCGCA	AGAAGCTTTA	AAATCATACA	TTGATGCTCA	AGATGATCTT	1380
	AAAGAGAAGG	AATCCCAAGC	GTATGCTGAT	GGTAAAATTT	CGGAGAGAAG	AGCAACGCGC	1440
30	TATACAAGAT	GCTCAAGCTA	AACTTGNAGA	GGCAAAACAA	AACGCAGGAN	CTTAAAGGCT	1500
	TAGANACGCT	GAAAAGANAG	CTAATGTTTA	TACAGACAAC	AAGGTCAAAG	AAAGCACAGA	1560
	TGCACAGAGG	AAAACATTGA	CTCGCTATGG	TTCTCAAATT	ATACACAATG	GTAAGGAAAT	1620
	CANATTAAAG	ACTACTAAAG	AAGAGTTTAA	TGCTTCTAAA	AGAACACTAT	CAAGAGTGTT	1680
35	AGCAGACATC	ACTGTAAATG	CTATGAAAAG	CATCTATTTA	AGGTATGACG	AAAATGGGGC	1740
	GATTACTTCA	CATACTATTG	ATAAAGATGG	CGTGAAAATT	AGTGGCGATA	AAGTTGATAT	1800
	AACAGCGAAT	AGAGAATTTA	ATGTATTCGC	AAATAATATT	AATAACAAAG	TTGGTAAAAA	1860
	TGACATTGTT	AATAGCCTAA	ACTTATCAAA	TGAAGGTCTT	GACATCAATG	TGAATAGAAT	1920
40	TGGTATTAAA	GGCGGAAATG	CTAACCGTTA	TGTACAAGTT	CAAAATGATT	TTATTGAACT	1980
	TGGCGGAATC	GTACAACGAA	CTTGGAAAGG	CAAACGATCA	ACCGATGATA	TATTCACACG	2040
	TCTTAAAGAT	GGACATCTAA	GGTTTAGAAA	TAATACCGCA	GGCGGGTCAC	TTTATATGTC	2100
	ACATTTTGGT	ATTTCAACAT	ATATTGATGG	AGAAGGCGAA	GACGGAGGTT	CATCCGGTAC	2160
45	TATTCAATGG	TGGGATAAAA	CTTACAGTGA	TAGCGGTATG	AATGGCATAA	CAATCAATTC	2220
	TTATGGCGGT	GTAGTCGCTT	TAACATCTGA	CTACAATCGA	ATTATTATCG	ATTCATATGC	2280
	TTCAGCTAAT	ATTGAAAGTA	GAGAAGCGCC	GATATATTTA	TCTCCGAACA	CCCCAAAATT	2340
	AAACCTGGNT	TTAANCCGAT	TCGCATTCAC	ATTATCAAAC	GCTGATAGGT	NCATTACGAA	2400
50	ACTGGCGGTT	ATATCATGTT	GGGTCAAGA	TGNAANCTAT	AAGTNCGGTG	CTGGATTAAAG	2460
	ATTTTCTAAA	CGTACCAATA	AAGGATTTGGT	TCAAGTCGTT	AATGGTGAAT	ATGCTACAGG	2520
	CGGAGACACT	ACAATTGAAT	CAGGTATGGC	CAAATTC AAC	TTAGTTAANC	GGAAGAGATG	2580
	GAAATAGTTA	CGTTAGCATT	CAAAGTTATG	ATTTATTGGC	GGTAGGTTCT	GATAATGCTG	2640
55	GCGATAGAGT	CGCTTCTAAT	TCTATTTATA	AGCGTACTTA	TTCAGCACCT	GCTAACTTAC	2700
	ACATTACTTC	TGCTGGAACA	ATTGGGCGTG	CTACTTCTGC	CAAAAAGTAT	AAAATTTCAA	2760

EP 0 841 394 A2

TCGAAAACCA ATACATCAAT GAAGACGATC AGTTCAGTCA TTCAAAAGAG ATTTTAAAGC 2820  
 TTCCAATTTCG TACATGGTTT GACAAATATG AATCGGAAAT AATGGCTAAA GAATTGGAAA 2880  
 5 GTGGTAAAAA GTTATCTGAT GATACTTTTA AACTTAGTCG ACATACTGGC TTAATAGCGG 2940  
 AAGAGGTTGA AGAATTAGGA TTTAATGAAT TTGTTATTTA TGATGACAAC GGAGAAATCG 3000  
 AAGGTATCGC ATACGATAGA CTTTGGGTTC ATCTANNACC TATTATTACC CNAAAACCAN 3060  
 CANTCAAANA NCGNAAAACN TANTGGNGGN NTTAACNTCT NATNGACNGC AACCAGGGTC 3120  
 10 CNCATNCCAC CTCCACATTA TACAANTCAC TTTCTCNCGT CACTA 3165

(2) INFORMATION FOR SEQ ID NO:109:

15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1229 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

AAAACAAGCG ATTATCACAG CAATTATTTT AATAGCACTG AGCTTGATTT TACCTGCATT 60  
 TGCAGTAGGT AATCACATTT CAATCTTCGT CATTTCTATG ATATTCTTTG CAGGTAGCTT 120  
 30 TGCTTTAATG TATGCACCTT TACTTAACGA AGCCATTAAA ACAATAGATC TTAATATGAC 180  
 AGGTGTGGCT ATTGGTTTTT ATAATTTAAT TATTAATGTG GCGGTATCTG TAGGTATTGC 240  
 GATTGCTGCG GCTCTAATCG ATTTTAAAGC ATTAAATTTT CCAGGCAATG ATGCATTAAG 300  
 TTCACATTTT GGTATTATTT TAATTATTTT AGGTTTAAATG AGTATTGGTC CGGATTAGTT 360  
 35 TTATTCGNC ATCTTTAAAT CCGTGCGGAC AACAATCTTG AAAAATAAAT AGATATTAAA 420  
 TCGCGAGATA TATCCGTATT TTATTGGTAA AATTAAATTA AAGAGATTAT ATTACACGAG 480  
 GAGTAGTAAG TATTGAAATT GGAGAAATAT ATAGATCACA CTTTATTGAA GCCTGAGTCA 540  
 ACACGTACGC AAATCGATCA AATCATCGAT GAAGCGAAAG CATACCATTT TAAATCTGTA 600  
 40 TGTGTGAATC CAACGCATGT TAAATATGCA GCAGAGCGAC TAGCTGATTC AGAGGTGCTC 660  
 GTTTGTACGG TAATAGGATT CCCATTAGGT GCGTCGACAA CTGCAACGAA AGCATTTGAA 720  
 ACAGAAGATG CAATTCAAAA TGGTGCAGAT GAAATTGACA TGGTCATCAA CATCGCGCA 780  
 TTAAAGATG GACGTTTTGA TGATGTACAA CAAGACATTG AAGCAGTGGT TAAAGCTGCG 840  
 45 AAAGGTCACA CAGTAAAAGT GATTATTGAG ACGGTATTGT TGGAACCATG ACGAAATTGT 900  
 AAAAGCGAGT GAATTAACAA AAGCGGCTGG TGCGGACTTC GTTAAACTT CAACAGGTTT 960  
 TGCAGGTGGC GGTGCGACTG CAGAAGACGT TAAATTAATG AAAGATACAA GTAGGTGCTG 1020  
 ATGTAGAAGT AAAAGCATCA GGTGGCGTAC GTAATTTAGA AGGATTTCAA TAAATGGTT 1080  
 50 GAAGCAGGTG CGACACGTAT TGGGCGCGAG CGCAGGCGTT CCAAATTATG CAAGGTTTAG 1140  
 AAGCAGATTC AAGATTACTA ATATATATAA AATTTGGGAG TGATAGCTAT GACAAAGACC 1200  
 ATTTAATCGN GTTCATTTTA TCCGTATGG 1229

55 (2) INFORMATION FOR SEQ ID NO:110:



# EP 0 841 394 A2

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1066 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

```

GGCACGAGAC GTAAATGNAA AGAAAACCGG CCTNCNTGAT TAAGANATTT AGAAAAGTGT      60
TTGTCATNAC GGTGTGCCCA GTGNTNAGTA TGTCCAAATA NAGGAAGTGA NAACGATGGN      120
ACCGTATTTT TAAAAANAGGC AGNACATANC ATTTTAGATA GGAAGTGATA ATATGAATTT      180
GGCTAAACGC ANATTACAAG GGAACNATT AACANAGAG ACTGTAGTGA AAATTTATGA      240
GGATACTAAT ATTGATACCT TAGNTTTATT AAATGAGGCG NACATCTTTA AGNANACATT      300
ATTTTGGTGC NATNAGTCAA ATTAACATG ATTTTAAATG CTAAAAGTGG NATATGTCCT      360
GAGAATTGTG GGCNCGAGAC TGTGGACAAT CACGAGATAT TAAACAANAA CAGCGATATG      420
CTTTAATTCC AGAGGAACAA ATTATCGATN GANCANAGGT GGCACATGAT AATCATATTG      480
GAACATATTG TATTGTTATG AGTGGTAGAG GACCGAGCGA TAAAGAAGTT GATCATATTA      540
GTAATCCTGT AAGAACGATT AAATCTCAAC ACCCGCAACT AAAAATCTGT GCATGTTTAG      600
GATTAACGGA TTGCCGACCA AGCTAAGAAA CTTAAGTCAG CTGTTGTAGA CAGATATANC      660
CACAAATATTA ATACCAAGTG AAAATTACCC ATTGATAACC GTCGTGCCAA CGCATAGTTA      720
TNAAGATAGA ACAGATACGA TAGAACTAAT GAAAGCGAAT AATATATCAC CATGTTCTGG      780
CGTGATTTGT GGTATGGGAG AATCTAATCA AGATATTGTT GATATGGCAT TTGCTTTAAA      840
AGAAATGGAT GCCGACAGTA TTCCGATTAA TTTTTCAT CCAATCAAAG GCACAAAGTT      900
TGGAAGCATG GATGATTAA CACCAATGAA ATGTTTAAAG ATCGTAGCAT TATTCCGATT      960
AATCAATCCT ACGAAAGAAA TTCGTATTGC TGGAGGAAGA GAGGTCAACT TACGTTCTGT      1020
ACAGCCATTA GCATTAAAAG CGGCGAANTN CAATATTTGG GCGGGG      1066

```

## (2) INFORMATION FOR SEQ ID NO:111:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3222 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

```

CCAAGTAAGT ATAGAAGCAG CATCCTTAGA AAAGAAAAAT GTAGATTGTA ACGAATTATT      60
ACAGCGTCTC AATGACGTTG AACAAACGGA ATACACCAGG TTCGCTTNAC CCTAAATTTT      120
TAAATGTTTC GTCACTTATT CAATATATTC AAGCAGCATA TCATGAACCT CTTAGAGAAG      180

```

EP 0 841 394 A2

	AATTTAAAAA	TTTAACACCT	TATGTGACGA	AATTATCGAA	AGTACATGGA	CCTAACCATC	240
	CATATTTAGT	CGAGTTAAAA	GAAACATATG	ATACATTTAA	AAGTGGCATG	TTAGAGCATA	300
5	TCCAAAAAGA	AGATGATGTT	GATTTTCAAA	ACTAATTAAA	TATGAACAAG	GTGAAGTAGT	360
	AAACGATATT	AATACAGTGA	TTGATGATTT	AGTATCTGAT	CACATTGCAA	CGGGACAATT	420
	GTTAGTGAAA	ATGAGCGATT	TAACATCTAG	CTATGAACCA	CCGATAGAGG	CATGTGGTAC	480
	GTGGCGACTC	GTTTATCAGA	GATTAAAAAGC	ACTTGAAGTG	TTAACACATG	AGCATGTTCA	540
10	TTTAGAGAAT	CATGTTTTAT	TTAAAAAAGT	ATCATAAATA	ACGCGATTAG	AAACTGTTGG	600
	CAAAGTAATG	TCAGCAGTTT	TTCGCTATAC	TTAACAGAAA	TTTAGTGATA	TGAACAGCAT	660
	TATTTGAAAA	GAAAAATGGT	CAACTTAGCA	TAAAAATTGA	TATGAAAATT	TAATGGTATA	720
	GATAATTAAT	TAGTAGCGTG	TTTTTTTAAT	AATTTATTCA	TGAATTTTAC	ATGCATATT	780
15	ATGATAAAAT	AAACATAATT	ATAATTCAC	GAGGTGCTAT	CGTGCTATCG	CTAACAAATG	840
	TATTACTTGA	GCGTGTAGGT	TTAATTATTA	TTNTGGGCCT	ATGTGTTGAT	GAATTATTCC	900
	ATATTTTTTAA	AAACTTTAAT	GAATCCGTCG	ACGCTACATG	GAAAGCACGT	TGGCAATTAT	960
	GTATTTTTTT	CAGTTTGT	GCCTTAATGT	TCTAATTTAA	CTGGTATCGT	CATCGATCAT	1020
20	CAACATAGTT	TGTCAGGAAG	TGTGTACTTC	CGTTTAGATG	ATGATGTATC	TTTAGCTAAC	1080
	ACACGTGTAT	TAACGATAGG	TGTCGCAGGA	TTAGTTGGTG	GCCCTTTTGT	AGGTCTATTT	1140
	GTTGGCGTTA	TTTCAGGTAT	TTTCAGAGTG	TATATGGGTG	GGGCGGATGC	ACAAGTTTAT	1200
	CTTATCTCAT	CTATATTTAT	CGGNATTAAT	TGCTGGTTAT	TTNNGGCTTA	CAAAGCTCAA	1260
25	AGGACGCAAA	GCGGTTACCC	GAGGTATTTG	GCGAAAAAGT	GGCCAATGAA	TTGGGANNTG	1320
	GTTATTGGGA	AATGAATCAA	ATGTNGGAGC	AATTTTAACA	TTTTTCCCCA	CGACAAAGCA	1380
	TATTGCGGTT	GACTCATATC	ATTAATTGCA	CTACCAATGA	TTATTGTTAA	TAGCGTTGGT	1440
	ACGGCGATTT	TTATGTCTAT	TATCATTTC	AACATTAAAA	GCAACAAGAG	CAAATGAAGC	1500
30	CTGTTCCTAA	CACATGACGT	ACTGCAATTG	ATGAACCAGA	CATTGCCGTA	TTTTAAAGAA	1560
	GGATTGAATA	GAGAATCGGC	ACAGCAAATT	GCGATGATTA	TTAAAAATTT	AATGAAAGTA	1620
	TCTGCCGTAG	CAATTACAAG	CAAAAATGAA	ATCTTATCGC	ATGTAGGTGC	AGGTAGTGAT	1680
	CATCACATAC	CAACAAATGA	AATATTAAAC	AGTCTGTCTA	AAGATGTATT	GAAATCAGGA	1740
35	AAGTTGAAAG	AAGTGCATAC	TAAAGAAGAG	ATTGGTTGTA	GTCATCCGAA	TTGCCCGCTT	1800
	AGAGCAGCTA	TCGTGACACC	ACTTGAGATG	CATGGTTCTA	TCGTCCGTAC	ATTGAAGATG	1860
	TATTTTACAA	ACCCTAATGA	TTTAACTTTT	GTGGAACGTC	AACTTGCAGA	AGGATTGGCA	1920
	AATATTTTTA	GTAGCCAAAT	TGAACTTGGT	GAAGCCGAAA	CGCAAAGTAA	GTTATTGAAA	1980
40	GATGCTGAGA	TTAAGTCATT	ACAGGCACAA	GTGAGTCCAC	ATTTTTTCTT	CAATTCAATT	2040
	AACACGATTT	CAGCTTTAGT	TAGAATAAAT	AGCGAAAAGG	CACGAGAGTT	ACTATTAGAA	2100
	TTGAGTTATT	TTTTCAGAGC	GAATTTACAA	GGCTCAAAGC	AACATACGGA	TTACTTTAGA	2160
	TAAAGAGTTA	AGTCAAGTGC	GTGCATACTT	ATCACTCGAA	CAAGCACNGT	TATCCAGGAA	2220
45	GATTTAATAT	CAATATTAAT	GTTGAAGACA	AATATCGCGA	TGTGCTTGTA	CCACCATTTT	2280
	TAATTCAAAT	TTTAGTTGAA	AATGCCATCA	AACATGCGTT	TACGAATCGA	AAGCAAGGTA	2340
	ACGATATTGA	CGTGTCAAGT	ATTAAAGAAA	CTGCAACACA	TGTACGTATT	ATTGTACAAG	2400
	ATAATGGTCA	GGGTATTTCT	AAAGATAAAA	TGCATTTGTT	GGGAGAAACA	TCTGTAGAAT	2460
50	CAGAACTCTG	AACTGGTAGT	GCTTTAGAAA	ATTTAAACTT	ACGCCTAAAN	GGATTATTTG	2520
	GAAAAATCCG	AGCATTACCA	ATGTGAATCC	GACATCGAGN	GGTACCACTT	TTTGGTGTGT	2580
	ACCTTCCTTA	TGAAAGACAA	GAGGAGGAAT	AAATATGAAA	GCATTAATCA	TAGATGATGA	2640
	GCCATTAGCA	CGCTAATGAA	TTAACATATT	TATTAAATGA	AATTGGTGGT	TTTGAAGAAA	2700
55	TTAATGAGGC	AGAAAAATGTA	AAAGAAACAT	TGGAAGCACC	TACTTGATCA	ATCAATATGA	2760
	CATTATATTT	TTAGATGTCA	ATTTAATGGA	TGAAAATGGG	ATCGAATTAG	GAGCTAAGAT	2820

# EP 0 841 394 A2

TCAAAAGATG AAAGAGCCAC CTGCGATTAT TTTTGCAACT GCACATGACC AATACGCAGT 2880  
 ACAGGCATTT GAATTAAATG CGACAGACTA TATTTTGAAA CCGTTTGGTC AAAAACGTAT 2940  
 5 TGAACAAGCA GTCAATAAAG TGCCTGCGAC TAAAGCCAAA GATGATAATA ACGCAAGTGC 3000  
 AATTGCGAAT GATATGTCGG CGAATTTTGA TCAAAGCTTA CCTGTTGAAA TTGACGATAA 3060  
 AATTCACATG TTAAAGCAAC ANAATATTAT TGGGATTGGC ACACATAATG GTATTACAAC 3120  
 CATACATACA ACGAATCATA AATACGAAAC AACAGAGCCA TTGAATCGTT ATGAAAAACG 3180  
 10 ATTGAATCCC ACTTATTTTA TACGTATTCA TCGTTCATAT AG 3222

## (2) INFORMATION FOR SEQ ID NO:112:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 805 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- 20 (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

ATCAATGTAA CCTAATAGTT TATGTCTATC TTGTGTACCA ACTACTACAT CGACACCAGG 60  
 AATTTCCATA ATTTTCAGCTG ATGAAGTTTG CGCATAACAA CCTGTACAC AGATTACAGC 120  
 30 ATCAGGATTG TGTCTTATTG CACGTCTAAT TATTTGACGA CTTTTTTTAT CACCTGTATT 180  
 CGTTACTGTA CAAGTATTAA TAACAAATAC ATCAGCATTG GCTTCAAAGT CAACGCGCTC 240  
 ATAGTTTGCT TCTTTAAATA ATTGCCAGAT TGCTTCAGTT TCATAATGGT TTACTTTTACA 300  
 ACCTAATGTG TGAACGCAA CTGTTGACAT AAATATTCAC CCCATTAATT CTTTTTCATA 360  
 35 ACTTATTGCA CTTAACGCAT ACAATGGCGC AGTTTCTGCC CGTAAAATTC TCGGTCCAAG 420  
 ACCAACAACCT GTACTAGTAT TACTAAATAA TGAAATTTCA ATTTCTGACA AACCACCCTC 480  
 AGGACCAAAA ATCATCAACA CTTTATCCTG AGCCTTGAAT TGTTGTAAAG TNTGCTTGAA 540  
 ATTGCTTAAC TCACCATCTT TTGCTTCCTC TTCATATGCA ATAAGAATAT AGTCNTAATT 600  
 40 ATCAATAGTA TCACAAATTA ATTTTAAATT CCGACTCCGA ATTGAATAGA TGGAATCCAC 660  
 TAAACGAATA GCTTTGTTCA GCAGCTTCTT TAATTATTTT TTGCCAACGC TCTATCTTTT 720  
 TGACAACTTT CGCCTCCGTT TAATTTAACA ATTGAGCGTT CCATGCTCAC AGCTATAAAT 780  
 GATGAAGCAC CCAATTCATA GCTTT 805

## (2) INFORMATION FOR SEQ ID NO:113:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1448 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- 55 (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

5  
GGCACGAGCG TAGATTAGC ATTTTCAAGA GACCAAGAAC ACAAAGAAAT ATGTACAGCA 60  
TCGTATAATG GAAGAAAGTG AACGTTTCAA TGAATGGATT GAGCAAGGCG CAGCAATCTA 120  
TATTTGTGGC GATGAAAAAT GTATGGCGAA AGATGTCCAT CAAGCCATTA AAGATGTATT 180  
10 GGTAAAAGAA CGTCATATTT CTCAAGAAGA AGCAGAGTTA TTATTGCGAC AAATGAAACA 240  
ACAACAACGC TATCAACGTG ATGTTTATTA GCGATTGGTG TTAAATATTT TAAGGTGTAA 300  
TGATGTAAAA AGATATAAAG GATGTTGCTC AACATGAATA TGCCATTAAT GATAGATTTA 360  
ACAAATAAAA ATGTCGTCAT AGTTGGNTGG AGGCGTCGTT GCAAGTCGTC CGGGCACCAA 420  
15 ACATTAAATC AATACGTTGA ACATATGGAC GGTCAATCAGT CCGACAATCA CTGAAAAACT 480  
TCAAATATG GTAGATAACG GTGTCGTCAA TATGGAAAGA AAAAGAATTT GAACCAAGCG 540  
ATATTGTAGA CGCGTATCTA GTTATTGCAG CAACCAATGA GCCACGTGTC AATGAAGCGG 600  
TAAACAAGC CTTACCTGAG CATGCCCTTT TTAATAATGT TGGAGATGCA TCAAATGCCA 660  
20 ATGTTGTATT TCCAAGTGCA CTACACCGCG ACAAGCTAAC TATCAGTGTA TCAACTGATG 720  
GTGCGAGTCC TAAGTTGACA AAATCAATTA TGGCAGAGCT TGAGGCGTTA TATCCACCAT 780  
CATACAGTTC GTATATCGAC TTTTATATA CTTGCCGACA GAAAATAAAA GTACTTGATA 840  
TAACATATAT ACGAAAAGCA ACAGTTACTG TCACAAATTG TGTCACAAGA ATATTTAAAT 900  
25 CATGACAAAC AAGCTCAATT TTTAGCGTGG TTGGATGTAA GATAATAATA GCGGACCGTC 960  
TAACCGTCTA AGGTAAGTCT TCTTATTTTA ACTTTAACGC TTAATCATTG AAATTAAGAC 1020  
ATGGGTGGCT TTGTGAATAG TCTAATAATG AAGGATTTAA GCGATAATGA TATGCGTTTT 1080  
AAATATGAAT ATTACAATAG AGAAAAAGAT ACGTAGAACA AACTTAATAA AATAGATGGA 1140  
30 TAAATTGAAA TCTGGTTGAA GTCGTTACTA TCATAGCGAC CTTTAGCCAG ATTTTTTGTG 1200  
CAATAGAAAG CAATAATAAA AATGATAGAT CAANATGAAA TACAGGACAG GATATACAAG 1260  
GATTAGTCAT GCCATGTTAT CAAGTAGGAA AATCAAACTT CACTATTGAT AGTTACGCAN 1320  
AAAGATTTTT TTGATAAAAT GAGATAACTT AAAGATAAAA AATTATATTA ATTATAATAT 1380  
35 TTAAGTTAAA GAGGGGATT ATGTAAATTG TATTTAAAGT GGAGGGAGAA AATAATATGA 1440  
ATAGTGAG 1448

## (2) INFORMATION FOR SEQ ID NO:114:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1127 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

CGACATGGAA GAACNTANAC CCTCGTNNAG GAGGCCCAGA AAGGNAGCNT NAAGNNAAAT 60  
55 TAAGNNAAGC GAAAAGCACC AACAATAAAA CCTGATGGCA GCATCATTCA ATGCGTGCCA 120  
CCAGGTTTTT ATGTTTTGTC AAGAAAATTA AATAAATCAT TAAATGATTC GGCCCATCGT 180

EP 0 841 394 A2

	AGGATGTGAA	TAAATACTAT	CTCGTAATAC	GGTATATGGA	ATGTTTTGAT	CAATCGCAAG	240
	TTTAATTATA	TTAATTAATT	CTTCAGATTG	CTTACCATAT	AATGTAGCAC	CTAAAATCAT	300
5	ATTATTTTCA	TTATTAATGA	CTACTTTAAA	TAAACCTCTT	GGATCATTGT	TAATTTTGTG	360
	ACGAGGTATA	GCACTTACTA	AAAGTTGATG	TTCAGTGTA	TCATAATGTT	GAGCGGCAGC	420
	TTCTTTACTA	GTTAATCCAA	CACGTGATA	TGGTGGATCT	ATAAATACTG	TATAAGGCAC	480
	GCTACCTCTA	TTGTCAGTCG	TACGTGACTG	ATTACCATAT	AACGCTGATT	TGATAATTCTG	540
10	ATAATCATCT	AAAGATATAT	ACGTAAATTG	AAGTCCGCCT	TTAACATCAC	CTGCAGCATA	600
	AATATGTGGC	ACAGATGTTT	GAAGTTGAGC	ATTGACTTTA	ATTTGCGCTC	TATCACCTAA	660
	TTGATATCA	GTATTTTCTA	AAGCTAAGGC	TGTATTCGGT	TTGCGCCCGA	TAGCCAAAAG	720
	TACTGCATCA	GCCTCANAGT	TGCCAACATT	AGNACGGA	GTTGTATGAT	GATCGTCAGA	780
15	TGACAAATCA	GTCGTTTCAA	CATTTGTATG	CAATGCAATG	CCTTTATTTT	CTAAGTCAGT	840
	AATAGCATGT	GCANCGACAT	CTTGATCTTC	GCCGTGGCAT	AAATGATTCA	CCACGTTCTA	900
	ATACNGTTAC	CTTACTACCT	AAATTCGCAA	ACATTGAAGC	AAATCTAAG	GCAATATAAC	960
	CGCCACCTAC	AATAACGAGA	TGCTTAGGTT	GATAGCTAAT	GTTTAATAAA	CCTGTTGAAT	1020
20	CGAAGACATG	TTTAGCTTGA	TCAAGGCCCT	TTATGTTAGG	AATGATAGAG	GTAGCACCAG	1080
	TATTAATAAT	GATATGAGGT	GCAGTAATAC	TAGCGACGAT	ATCGCCG		1127

(2) INFORMATION FOR SEQ ID NO:115:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 876 base pairs

(B) TYPE: nucleic acid

30

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

	GAAGANATAN	CGTGGTNTTG	AAGATTGTTGG	ANGTTAATAG	ATNAAGGGGA	ATATGCANTT	60
40	ATTACCGGGA	GNNNAGTATC	CAAGNAAATC	TAGAGTCAAT	AGGTTAATAA	TCTTATGCCT	120
	TTTTAATTTT	GAATAAAAGT	GGGGTGGTGT	AATGTTTGGA	TTTACCCANC	CGACCCGAAC	180
	AAGATTGNCG	TTTNACGCGA	TTAGAAGNAA	ATGATAAGCC	TATGTTTGNN	AAATTCGACA	240
	GAATAGAAGA	CAGTCTGAGA	ACGCAAGAAA	AAATTTATGA	CAAGTTAGAT	AGAAATTTCTG	300
45	AAGAACTAAG	GCACGAGCGA	CAAGGTAGAA	GATGAAAAGA	ATAAAGAAAA	GAATGCCAAA	360
	AATATTAGAG	ACATAAAAAAT	GTGGATTCTA	GGATTAATAG	GGACGATCTT	CTAAGTACGA	420
	TTTGTCATAG	CCTTACTAAG	AACTATTTTTT	GGTATTTAAA	GGAGGTGATT	ACCATGCTTA	480
	AAGGGATTTT	AGGATATAGC	TTCTGGGCGT	GCTTCTGGTT	TGGTAAATGT	AAATAACAGT	540
50	TAAGAGTCAG	TGCTTCGGCA	CTGGCTTTTTT	ATTTTGATTG	AAATGAGGTG	CATACATGGG	600
	ATTACCTAAT	CCGAAAAATA	GAAAGCCAC	AGCTAGTGAA	GTGGTTGAAT	GGGCGTTATA	660
	TATCGCTAAA	AACAAAATAG	CTATTGATGT	ACCTGGTTCT	GGAATGGGAG	CACAATGCTG	720
	GGAAATTTACC	TAATTATTTA	TCCGATAAAA	TATTGGGGAT	TTAGAACATG	GGGAAATGCT	780
55	GATGCTATGG	CTCAGAAATC	CAATTAATAG	AGGTAGAGAT	TTCAAGATAA	TTAGAATCAA	840
	AAGACTTGGT	ACACAACCAG	GCGACTGGGT	TTTTGG			876

## (2) INFORMATION FOR SEQ ID NO:116:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1789 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

```

GGCACGAGCG GCACGAGCGT GTTGTATCAA GATTTTGTAG GCAGTTTAC AACGTCCGAT   60
TCAGCAAGTT ATGCACAAGA TTTTAAATCT GAGGAAAACG CTAAAAAGAT TGCTGAAACT   120
TTAAATCTTT TATATCAATT AACAGGCAAT CAAAACGGTG TGAAAGTTGT GAAAGAAAGTT   180
GTGGATAGAA CTGACTTGTC ATCTGATAAA TCAGTTGATA GCGAAACAAT GTAACATATAC   240
TAAGTTATGA GCATTACGCT CATAGCTTTC TTAGAAAGTA GGTGTAGTTT TGGATGATAT   300
TCAGAAAATA AAAAAAGAGC TTTCTGAATT AGTTGAACGT GTTGATGATG TTGAAATACT   360
AGCAAAACGAA ACAGCTGATC ATGTGCTTGA ACTTAGAGAG GAACATAAGC AACATCATAA   420
TGAACTAAGA GAATCTCATA AAGAACTTAA AGATAAGCAA GATAAAGTTG TAGATGAGAA   480
TTTAGAGCAA ACAAGATAT TAAACAGAAT TGAAGAAAGA TATCANACGC AAGTAGNTGT   540
TGNGCAAAAA AATGAAGAAA AGACACTCGC CCAAAATAAA TGGCTCGTAG GTGCCATATG   600
GGCGCTTGTA ACAATTGTTA TGATTGCAGT CATTACTGCA TCAATTNCTG CGTTATTACC   660
TTAAGGGAGG TGGACATAAT GAGTTGGGCA AGATGGTTAT CATGTTATTT GTNTGGTCGT   720
AAATGTAAT AATGTTTTTG GTCAGTGCAT CGGCACTGGC TTTTATTTTT GATTGAAAAG   780
AGGTACGTAC ATGGTATTAC ACAGCTCACA AGACAGGAAG CATACTCCAA GTGAAGTTGG   840
GAAGTGTTGT TAATACCAAG TAAGTAGGAT ATCTGANATG TATAATAGAG TAAAAATGAA   900
ATCTTTTAT TATAGACACA TATAAAAAGT GTATAGTAAT ATATGTATGT ATAATTAAAT   960
GATAATCATT TCATAATTAT TGTATATAAC TAAATAACTA CTTAACANAA ATAATTATGC  1020
TTTAGAGNTG ACCANNATGA NNNANNCCAG CATTTACATT ACTTTTATTC ATTGCCCTNA  1080
CGTTGACNAC AAGTCCCANT TGTAATGGT AGCGAGAAAA GCGNAGNAAT AAATGCGAAA  1140
GATTTGCGAA AAAAGTCTGA ATTCAGGGN ACAGCTTTAG NCAATCTTAN NCANATCTAT  1200
TATTACNATG NNANAGCTAN AACTGAAAAT AAAGAGAGTC CNCGACCACA TTTTACAGC  1260
ATACTATATT GTTTANAGGC TTTTACAG ATCATTCGTG GTATANCGAT TTATTAGTAG  1320
ATTNTGATTC NNAGGATATT GTTNATAAAA ATAAAGGGNA AANAGTAGAC TTGTATGGTG  1380
CTTATTATGG TTATCAATGT GCGGGTGGTA CACCACAAA AACAGCTTGT ATGTATGGTG  1440
GTGTAACGTT ACATGATAAT AATCGATTGA CCGAAGAGAA AAAAGTGCCG ATCAATTTAT  1500
GGCTAGACGG TAAACANAAT ACAGTACCTT TGGAAACGGT TAAAACGAAT AAGAAAAATG  1560
TAACTGTTCA GGAGTTGGAT CTTCAAGCAA GACGTTATTT ACAGGAAAAA TATAATTTAT  1620
ATAACTCTGA TGTTTTTGAT GGGAAAGGTC AGAGGGGATT AATCGTGTTT CATACTTCTA  1680
CAGAACCTTC GGTAAATTAC GATTAATTTG GTGCTCAAGG ACAGTATTCA NATACACTAT  1740
TAAGAATNTA TAGAGATAAT AAAACGATTA ACTCTGAAA CNTGCGTAG   1789

```

EP 0 841 394 A2

(2) INFORMATION FOR SEQ ID NO:117:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1410 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

CGATCTACAT	TCGCTGCGCC	AATACACACA	ATAGGATAAT	CTTCATTAA	AACATATGCC	60
TTTCCCATAA	CATATTCCTT	TTGTATTAAT	CCTGAAATAA	TGTTTGCTAC	GCTGGGTCTA	120
GATAATCCAA	TTGCCTCAGC	AAGTTCACGT	TGTGAAATAA	NCGGATTATC	TTTAATTCTT	180
TTTAAATTTT	CTTCTCAGN	ATCGCTCATA	ACCATCCCTC	CTCATATTGA	CTTAAATATC	240
TTGAATCAGT	GAATTCACCA	AAAATATCGT	AGCATGCTGT	TATTTCTGTT	GTAAACGCTT	300
ACATTTTNCC	GTTATATCAA	TGATATTGT	TTATAAAATA	CGCATAATTT	CGTAAAAAGT	360
AAGTTCATTG	CCATCGTGTA	AAAGNTACTC	GAGCAAATGA	TAAATGTTTA	TACAGTATGA	420
TATTTGTCGA	CAATAAATTA	TGACATTTT	ATGTCTTATA	TTTTTCTATA	AAAAAACGTG	480
ACTCAATTATC	GATTTTTTCA	TCGACTGAAT	CACGTTACGA	TGTTTTATTT	GTTTTTNCTA	540
TTCAATATGA	TGATGCTATG	CGTCCAACTG	CACCTCGATT	AATCATGTCT	TGCTTAGTAA	600
CAATTGGATA	TTTTTTAAAT	ACGGATTGAA	CCCAAGATAA	CCCTGCAAGC	GCACCCTACA	660
ATCGCCATTA	CTACACCATA	AATTACAATA	GTCATTGGAT	GATTAAATCC	ANACATTACT	720
AAAAATCCTN	CAATCGGTGT	AGCTGTACCT	GTCGCATCGT	TAATTAAACC	TGACATAGCA	780
ATAATTAAAC	CAGCAATCGC	ACCACCAAAG	AAATTGTGTA	TATAGATTGG	AATTGGATTG	840
GCTGATACAA	TATCTGCTTG	TGATAAAGGT	TCAATACTTA	CTGCAATCGT	AGACTTACGA	900
TCACCTAATT	TTAAGCGATG	GAATAGCGTC	CCATTCATAA	ATGCCGAACT	AAATGCTGCC	960
ATGGCACCAA	TAGCCATAGG	TACACCGT	AAACCTAATA	ATGCTGTCAA	TGCCATTGAA	1020
CTCAATGGCG	CTGTGCCGAC	AACCGTGATA	ATACCACCTA	AAATGATACC	CATGATAATT	1080
GGATTCGTAT	TCGTACTACT	TTGGATAATA	TCACCAATTC	GAATCAATGT	ACTATTCACT	1140
ACTGGCGTTA	ATAATACTGC	AATAAGACGA	GCTAATGGCG	CTAAGATGAC	AATCGGATCC	1200
AATTAAGTCG	ACAACCATCC	GGGCACATAT	TTCTCTTGTA	TATTTTCATCA	CATATCCAAT	1260
CAATATTTAA	CCCGNAAAGG	AAACCTGGGT	AATAAATCCA	TACCACCGCA	ACGCTCGGGC	1320
CGGTTATTAA	CTTAATGGAT	ATTCTGGGGG	ATACAACCGA	TTAGGTTAAT	GCTTGTGAA	1380
TCCCAGGCGG	GAGGGGACAA	CAACCTAGGA				1410

(2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 926 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

EP 0 841 394 A2

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

AATTATCCG	GTTTCGCCGT	CTTTGTTTGT	GGCTATGTTA	CATTCAACAA	TAGATTTGCC	60
AGTGATACTG	TCATCTTCGT	CACGGTTATA	ATAATCATCA	CGGTAAAGTA	GCATTGCTAA	120
ACTTGCAATCG	GCTTCTATTC	CGCCTGATTC	TTTCATGTCT	GATAGCATTG	GTCGTTTATC	180
CTGCCTAGAT	TCAACACCAC	GATTCAGTTG	TGAAAGTAGT	ACGATGATTG	CGCCTGTCTC	240
GTTAGCGATT	ATCTTTAAGT	CACGTGATAT	CTTTTCTACT	GCTACACGTC	TATCAACTTT	300
CGCATCAGTA	TCCATCAGTT	GAAGATAATC	TATAAAAATA	ACTTGGTTCG	TCGGTCTGAA	360
TGCCTCATTG	GTTTGAGCTC	GCACATCTTG	CGGTGGTAAT	ATTTACTTTT	GGTCAAGAAA	420
TATCAAATAC	CTAACTTCAA	TGATTGGAT	CCCATTCGAT	TTGTTAACTT	CGTTAAGTCA	480
TCCGGTGTTA	AGTTCCTGAT	TTCTTTTATC	TTGGTTAACT	CAATACCAGT	AATTGTTGAT	540
AACATACGTT	TCAATACTGA	TGTGCCAGTT	GTTTCAAGAC	TAAAGAAAGA	TGTTTTATAT	600
CCATTTTCGTG	CTATGTTCAA	CATCATGTTT	AATGCAAAAC	CTGTTTACC	CACTGAGGGA	660
CGCGCGTGCC	ATGACGATTA	ATTGCGACGG	CTCTAATCCC	CCTATTTTGT	AATCCATGAG	720
CTTATAACCC	GTCTTAATTT	GCTTCTTAGG	GCTATCGCTG	TATAACTCAT	CAACAACTC	780
CTCAACAAAC	TTCTTGGTTC	CATCTTCTTT	TTTGTTAGTA	ATCGTTTTTA	AATCCTTGAG	840
TTCATCAATC	AAGTTGTAA	AGTTTGGTT	CGTAGGTTGT	TGTTTAAACT	CAGTTACCAA	900
TTCGTTAGCT	TTGTTGAGCT	CGTGCC				926

(2) INFORMATION FOR SEQ ID NO:119:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1406 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

CCAATAATCC	ATTNCAAGGC	TTTTTCGTTG	CCCCAATNAA	GAAGCACGAG	ACCCATCCCG	60
GCTATTAAAT	ACCTGCAATT	TGAAATTTGC	CAAGGGCACC	ATGGTGTGAC	CGACAAGTAT	120
TAAATCCTG	CACCAGCGAA	AAGCATTACC	TAATGAATTA	TTATCATTAA	TCGATATTAT	180
TGTGCCAAAC	GAAACAGAAG	CCGAATTGTT	ATCTGGGATT	AAAGTAACTA	ATGAACAATC	240
TATGAAAGAC	AATGCCAATT	ACTTTTATC	TTTAGGCATT	AAGACTGTTT	TGATTACGCT	300
AGGTAAGCAA	GGTACATATT	TTGCTACTAA	AAATCAAAGC	CAACACATCG	AAGCTTATAA	360
AGTAAATGCG	ATTGATACAA	CTGCTGCAGG	CGACACATTT	ATTGGTGCAT	TTGTCAGTCG	420
CTTAAACAAG	TCGCAAGATA	ACTTAGCAGA	TGCTATTGAT	TTTGTAATA	AAGCGAGCTC	480
ACTCACTGTA	CAAAAAACAC	GGCGCGCAAG	CATCTATTCC	TCTACTAGAA	GAAGTAAATC	540
AAGTTTAAAT	GAATCAAACA	CAGCTATGAT	ATGAAGGTTT	AGCATATAAC	ATGCAACATT	600



# EP 0 841 394 A2

	CGTATATCAT	GGCTGTGCTT	TTTTATCTTT	ATAAAACATC	ATCTATTAGA	AATAATTTTC	660
	CACTAAACCT	ATGCTTGTTG	ACTCATGTTT	AGTTATAAAT	GAAGTGACAA	TTTTTTGTAA	720
5	TCTTTTAAAC	TTCCAAATTA	TTGCATATAA	ATATGCTATA	TTAATGATAA	TAATTATCAA	780
	TTAAAAGGAG	GTTATGCTAT	GTCTAAAGAA	GCTGGTCATA	CATTTTTAGC	TAAATTAGGA	840
	AAAACCTCGC	TACGCCCCGG	TGGGTAAAGA	AGCAACAGAT	TGGNTTAATA	CANCAAGGGG	900
	CATTTTCACA	AGATAAACAA	GTGTTAGAAG	TGGCATTGTA	ATATTGTGCA	CAACATCTAT	960
10	TTATCTAGCT	CATACATATG	GCTGTCACAT	TCAAGGCGTT	GGATATAAAT	AAGAAAGCAT	1020
	TAGAAAAAGC	ACAGGAAAAT	TTCCAGCAGC	AGGTCTTGAA	TCATATATTC	AAGTTCAACA	1080
	AGCGAATTGC	TGTTAAATTG	CCCTTTGATG	ACAATCAATT	CNATGTCGTT	TTAAATGAAG	1140
	CAATGTTAAC	AATGTTACCC	ATCGCCATAA	AGGAAAAAGC	ATTACGCGAG	TNCTACCGAG	1200
15	TCATAAAGCC	NGGGGGTNTN	TTGTTAACCA	CATGATNTTG	TCATCGTTAA	TGAATCANAT	1260
	NCCACANATG	TTGTAAATC	ATTATCTGCA	GCAATTAATG	TCNATGTCTC	ACCGCAGACG	1320
	AAACTTGGCT	GGTTAGATTT	ATATCATCAA	AGCTGGTTTT	AATCAGGTGC	ATTATCCATA	1380
	CTGGTCCATG	AGTTTAATGA	CACCAA				1406

## (2) INFORMATION FOR SEQ ID NO:120:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 654 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

35	GGACGTAATN	ATACGCCTTT	TCTGATTATT	TCTTTACTAT	TCTAAAAACT	GTTGTTCTAC	60
	TAATACCTAC	CATTTTGAAA	TTTCATTGG	CGTGACATT	CTACTGCGA	ATAATTGTTT	120
	AATATATTTA	CGTTTATCTT	CAGAAAGTTG	AGGTCGTCCT	CCATACCTAC	CACGTTCTTT	180
40	TGCAGATTCT	AGACCATTG	TTGTTCTTTC	TTTAAATAAT	ATATTTCTG	AATCTGCTAA	240
	CATTTTAAAT	ACTTCAATAA	ATGATTGTCC	TTCTTCACTA	TTAGTATTTA	TGCCTAAATC	300
	AAGCACTTCT	AAATGGATAT	CATATGGTAA	TATCTACTT	TCAAAAAATT	CCATGAGTTT	360
	TTTGGGTGTG	AATCCTAGTA	CTTTTAATTT	TTCTATCACT	AATGTATCAT	CTGTTTTCAA	420
45	CTCGTCTAAA	CAATGCTCAA	GTTCTAGTTG	TTCTTTATAA	TAAACATTTA	CTTGCTCAGA	480
	ATAAACTAAG	TCACAACCAT	TTATTAAGAG	TTTCGTTAAT	TGAGTATCTA	GATTTTGATC	540
	AGGATATGCT	ACACGTGCGT	AACCAATTTT	CCTCATTTTA	TACACCCTTC	CAGCTTATTC	600
	ATTTTCTCCG	ATTAATAAAC	CTTCGCTCGT	GCCGTTTCAA	CAAGTCGGGC	CCTG	654

## (2) INFORMATION FOR SEQ ID NO:121:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1213 base pairs
- (B) TYPE: nucleic acid

EP 0 841 394 A2

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

10	GCAAAGGCTT TGAAGATATG AGTGGTAAAG AATGTACCTA AGGATGACAT TGAATAGAGG	60
	GAGTAAAGAA AGCGACAGCT GATGCATCCT CGCTAGTTTG GACGGCATGA ATAGAATGAC	120
	GTTAGGAAAT GAAGTTCTGA CTAACAAAAA GATTGAACT GAAATCAAGA TGCTTGAGAA	180
	AAAAATTGAC CAAATGGATA AATCAGAAAA TAATTCACAA GAAGCAGAAG TTGCTAAAGC	240
15	ACTTATTAAG TTAGCGGGTG TTAATAATGA TTAATGAAAT GTTAAACCCG AAACAACAAG	300
	AAGTCTGGAC CTGCTTTTATA AACGATAAAC CCAAAGTATT AATAGCGAGT GGTGCCAAAA	360
	GGGCAGGTAA AACATATGTG TTCATCCTGC TTTTTTTAAT GCATATAGCT ACTTATAAAG	420
	ACAAGGGGCT TAACCTTCATT ATTGGAGGAG CAACACAAGC ATCTATAAGA CGTAACATAC	480
20	TAGATGATAT GGAGTTAATA CTAGGTAGAG AGTTAACT CGACAAATCT AACGCAGTCA	540
	AAATATTCGG TAATAAAGTG TATGTATTTCG ACGGACAAAA CTCGGATGCA TGGAAAAAAG	600
	CGCGTGGTTT TACTTCAGCA GGTGCTTTT TAAATGAGGG AACAGCATT AACAATATGT	660
	TTATTAAAGA AGTGTCTCA CGTTGTAGTT ACAAAGGCGC GAGAATATTA ATTGATACCA	720
25	ACCCCGAAAA CCCGATGCAT CCAGTTAAAA AAGATTACAT TGATAAGAGT GGTCAACGGT	780
	TATCGAATGG AAGACTAAAT ATCAAAGCAT TCCAATTAC TTTGTTTGAC AATACATTTT	840
	TAGATGAAGA ATATATTGAA TCGATTATCG CAAGTACACC AACAGGAATG TTCACAGATC	900
	GTGATATTTA TGGTAAGTGG GTTCTGCTG AAGGTGTTGT ATATAAAGAT TTCAAAGAAA	960
30	AAGTTCATTA CATCACAGAA GAAGAATTTA AACTAAACA AATAAAAAGG AAATATGCAG	1020
	GCGTCGACTG GGGATATGAG CATTATGGTT CTATTATGGT TGTAGCGGAA GACTTCGACG	1080
	GAAACAAGTA CGTTATTGAA GAACACGCAC ACAGACATAA AGAAATAGAT GACTGGGTAG	1140
	CTATTGCAAA AGAGTTATAA AAGCATGGCG ATATCCTTTT TATGTGATAC GCTAGACCTG	1200
35	ACATATTGAC ATT	1213

(2) INFORMATION FOR SEQ ID NO:122:

(i) SEQUENCE CHARACTERISTICS:

40

(A) LENGTH: 1141 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

45

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

50

	CATTATTATT TGTCGTTAAA TACAATAGAA AATACTATAC CTGTATATGC AATTCGACAA	60
	TAGATAAATT ATTAAATATG CTTACAACAA TCTTAATATC CTTTAACGCA CTACAATAGT	120
55	GCTCTGATAA TAGGTTATAA ATGTACGTAA AACCATTGTC TCAATAAAAA TGAAAACGTA	180
	TACTTCAAGA AGGATGGGTT ACTTAATATA AACAAGGGGG TAACATATAT GACTTTATAT	240

EP 0 841 394 A2

	TTAGATGGTG	ANACACTAAC	AATTGAGGAT	ATTAACATCAT	TTTCACAACA	ACACTCAAAG	300
	ATTGANATTA	TTGATGATGC	GTTAGAACGT	NTCAAAAAA	GTAGAGCGGG	AGGTGAACGT	360
	ATTATTGAAA	ATGANGAAAC	GGGNTTGACG	GGTATCACTA	CAGGGTTGNG	GGGTTTATTT	420
5	AGTGATGTAC	GTATTAGCAC	CCGACGCAAT	ATNATGAATN	ACAAGTGAAT	CTGATACGCT	480
	CACATGCCTG	TGGACTAGGT	GAGCCATTTC	CAAAAGAAGT	AGCATTAGTC	ATGATGATTT	540
	NACGATTGAA	TACATTATTA	AAAGGTCATT	CAGGTGCCAC	TTAGTAATT	AGTGAGACAA	600
	TTACAATTTT	TTATAAATGA	ACGTATTATA	CCGATAATCC	CACAACAAGG	CTCTCTCGGT	660
10	GCATCAGGAG	ATTTAGCGCC	ATTATCACAT	TTAGCATTAG	CATTAATTGG	TGAAGGGAAG	720
	TATTGTACAG	AGGGGAAGAA	AAGGATAGTG	ACGATGTATT	AAGAGAATTA	AATAGACAAC	780
	CTTTGAACTT	TCAGGCTAAA	GAGGGTTTAG	CATTGATTAA	TGGTACGCAA	GCTATGACAG	840
	CTCAAGGTGT	CATTAGTTAT	ATAGAAGCAG	AAGATTTAGG	TTACCAATCT	GAATGGATTG	900
15	CTGCATTAAAC	GCATCAGTCT	CTTAATGGCA	TTATAGATGC	ATATCGACAT	GATGTGCACT	960
	CTGTTCTGTA	TTTTCAAGAA	CAGATTAATG	TGGCAGCGCG	TATGCGTGAT	TGGTTAGAAG	1020
	GATCAACATT	AACGACGCGA	CAAGCAGAAA	TACGTGTACA	AGATGCATAT	ACGTTGCGTT	1080
	GTATACCACA	AATCCATGGC	GCTGGTGCCC	AAGTATTCAA	TTATGTTAAA	CAGCAATTAG	1140
20	A						1141

(2) INFORMATION FOR SEQ ID NO:123:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1258 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

	TTTGTGCTTC	GTCAATCTTG	TCACCANCTA	ACTCTAAACG	CTTTCCTTTT	GTCGTTTTTCG	60
	GGNTCCCACC	ACAATTAATC	TGTTTTGATT	TCTGACATAT	CAATCTTCCT	AGTTAACCCA	120
40	AGCCAAATCT	TTACCATACT	CAATTTCTTT	ACCTTGGAAC	AAGCTGTTTT	TCTTGTTGAG	180
	TACCACATAT	CTACCTTTGA	CGGTATTAGA	ACTAAGCTCT	ATATAAAAAT	CCAATGCCAT	240
	TTTATAGGTT	GTACATAATT	GCTTTAAAAC	TTCATATCTA	GTNTGATAAG	AAGTCCATGA	300
	CGTAGTACGT	AAGCCATCGT	ATTCGGTNTG	TTCAGAAACT	TCCCAACCTG	TATCGCTCAA	360
45	CACATCTTTC	AATGCTTCTG	AAGTTGTCTT	TTTCTCAAAT	TTGCCTGGTG	CATACGGTTT	420
	AGCTGTTGTT	ATATCAGCAA	GATAAGACGC	TATACATTCT	ATCTCTGTGT	AGCCGTCCAT	480
	CGTATCTTGA	ACCCAGTTAA	TAATAAATTC	ACGCCATTGT	TTGTTTGAAT	CCCTTATAAT	540
50	AACACGATGT	CGTTCACGGA	ACNTTTCAGC	TCTTTCCTGAT	GATATGAGCA	GTTCAAGCAT	600
	TTCTGAATTG	TCATTAACAT	TACGTTTATG	AATCGCTCTA	ACTAAGGAAG	GATCATCAGT	660
	AGAAAGGAAA	TCTATAATCT	TGTCGTTAAA	ATCTAAAACA	TGTATCACAC	TCTCATCTCC	720
	TTTCTATAAA	TATCTATCTT	GCCATTTAAC	CGTCGTATCA	AAGACGTTTT	CAGGTTGTAT	780
55	GATTAATTCA	CTGTACCCAG	AATCAACATT	GAAATAATTA	CTTCCAAACG	ATTTCTCGCT	840
	CCAACATTGG	TTCCTCATTT	GATGACAACA	CTTNNNGCTT	GNATATCTAT	TTTCACTAAA	900

EP 0 841 394 A2

	TCACCTTTTT	GTATAATGAC	ATCCCTTGCG	CCTTTCGGNT	TCGGTAGAAT	CTCCGTATTG	960
	AATGAACCTA	ATCCATTTCAT	CTCCATCCAC	TTATAACCGT	NATACTTCGC	ACTATAGATA	1020
	GCTATGATAG	AAGCTGGACG	CTGATAAAAC	TTACCGCCAT	CTATCCACTC	TTTCTCATCC	1080
5	ATATCAATAG	GTTTACGTCT	ATCTGGGTCT	TTAATGTGAT	CAAATTTCCA	AGTTTTAATA	1140
	GAAAATTTAT	TACCTACTCT	TCTGAGCCGC	ATATAAACAN	CGATTCTGTC	CAAGTTATAC	1200
	ATTATCGGTT	TATTCTGATA	GTCGTATATC	TTTTTGGGGT	CTCCTTTTTG	GTTATACG	1258

(2) INFORMATION FOR SEQ ID NO:124:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1620 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:

25	TCGAATTTGG	TGTTGNCCGT	TCCTGTATNA	CATTCTGNCT	TATCTAATGG	GGGACGTTAT	60
	GATGAGTGCC	CACNNATTAG	GGATGATCGT	NCGAGAGTAA	GTGTTGGTGC	AAGGTCAAAGT	120
	GTGTTGCGAC	CTTTCCACAA	ANNTAGGGTT	AATCATCATT	GATGAAGAAC	ATGAATCTAC	180
	ATATAAACAA	GAAGATTATC	CGAGATATCA	CGCTAGAGAA	ATTGCCCAAT	GGCGAAGTGA	240
30	ATATCATCAC	TGTCCAGTCA	TTTTAGGAAG	TGCAACACCA	TGTCTTGAAA	GTTATGCACG	300
	AGCTGAAAAA	GACGTTTATC	ATTTGCTATC	ATTACCAAAC	AGAGTGAACC	AACAAGCTTT	360
	ACCTGAAATT	GATATAGTAG	ACATGCGTGA	AGAATTGAGT	GAAGGTAATC	GGTCAATGTT	420
	TTCAAAAGAT	TTACGTGAAG	CCATACAATT	AAGATNAAGA	TCGACAGGAA	CAAGTTGTTT	480
35	TATTTTAA	TCCGACGTGG	TTATGCATCG	TTTATGTTAT	GTCCGGATTG	TGGATATGTA	540
	CCGCAATGTC	CAANCTGTGA	TATTTTCATTA	ACGTATCATA	AAACGACAGA	CTTATTAAAA	600
	TGTCACTATT	GTGGTTACCA	AGAGACGCCA	CCGAATCAAT	GTCCAAATTG	TGAGAGTGAA	660
40	CACATTTCGAC	AAGTAGGTAC	TGGTACTCAG	AAAGTTGAAG	AACTATTGCA	ACAAGAATTT	720
	GAAGATGCGC	GCATAATTAG	GATGGATGTA	GATACAACCT	CAAAGAAAGG	TGCACATGAA	780
	AAGTTATTGA	CTGAATTCGA	AAAAGGTAAC	GGTGACATTT	TACTAGGTAC	TCAGATGATT	840
	GCGAAAGGAT	TAGATTATCC	AAATATTACT	TTAGTTGGTG	TGCTGAATGC	NGATACAATG	900
45	TTAAATTTAC	CTGATTTTCG	GGCGAGCGAA	CGTACTTATC	AACTATTAAAC	GCAAGTGGCT	960
	GGTAGAGCTG	GTCGTCATGA	AAAGGCAGGT	CAAGTCATCA	TTCAAACGTA	TAATCCAGAT	1020
	CATTATTCAA	TATTGGATTG	TTCAAAAAAA	TGATTATTTA	ACATTTTATC	GTCAGGAAAT	1080
	GGAATATCGT	CAAATTAGGA	AAGTATCCAC	CGTATTTATT	ATTTGTTTAN	TTTCACAATC	1140
50	TCACATAAAG	AAATGAAGAA	GGTTATGGAA	GCATCGCAGC	ATGTTTCATAA	AATTTTATTA	1200
	CAGCATTTAC	AAGAAAAAGC	GCTTGGTACT	AGGTCCATCC	TCCGGCAGCA	CTTGCGAGAA	1260
	TCAACAATGA	ATTTAGATTC	CCAAATTTTA	GTGAAATATA	AAAGTGAACC	TGGATTATTA	1320
	CAAGCCANTC	AGTTTTTAGA	TGACTATTAC	CATGAAAAAT	TTATAAAAGA	AAAATTAGCA	1380
55	TTGAAGATTG	ATATTGGANC	CACAGATGAT	GATGTAACAT	TACTAATTAT	TAGAAACAAG	1440
	NNCAAGTATT	GTACGAGTAT	TTGAACCCAG	TGTGTGAATA	TTTACTTNAN	GNACAAGAAA	1500

# EP 0 841 394 A2

AAGGGCAGAA TATACAACCTG TTAACCTATTT AAATTAGCAG TTTATATTCT GTCTTTTAT 1560  
ATGGCTTTAT AACTTACGTG ATTTTGGTTT GATAAGGAAT TTATTAGTAT TTTCATTTAC 1620

5 (2) INFORMATION FOR SEQ ID NO:125:

(i) SEQUENCE CHARACTERISTICS:

10 (A) LENGTH: 875 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

20 AATTGGGCAC GGGTANACAG TTGTAGCGTT ATTAAGNATT TTAACCTTTG CGAAGTTAGC 60  
AACAATTTNC CAATGATGTT ATTTTATCAA TGTTACCTCA AGCAGCAACT ACAANCGATT 120  
GCGTTACCAG TATCAGCTGG TATCGGTGGT ATAAAAGAAT TAACATCATT AGCAGNTATT 180  
TTAAATGGTG TCATTATTTA TGCCCTAGGT AATAAATTCN TGAAGCTTTT CCGAATTACT 240  
25 AACCCCTATTG CCCGAGGATT AGCACTTGGA ACNAGTGGTC ACACATTAGG TGTAGCACCA 300  
GCCAAAGAAT TAGGACCTGT AGAAGAATCA ATGGCAAGTA TAGCTTTAGT GTTAGTTGGT 360  
GTAGTTGTTG TAGCAGTTGT GCCTGTCCTTT GTAGCAATAT TCTTCTAAAA CGAAAAACCT 420  
AAGCAAGATA ATAGCAATTT GAGCCATTGT TATTATCGTA AAAAAACGTC TATACTCCAG 480  
30 TTTATAACTG GGATATAGAC GTTTTATGT ATTTATTACT TTTTACTAGG AATATAAAAC 540  
TGTGCATGNC GATAATGAAA TACGATGTCA GACGAATCAA AGGGTTTGCC AGTCATTGTA 600  
TAAAAAGTCT GGTGGTAACG TAAACATGGT TCACCTGTAG ACAATTGTAG TAATGAAGCT 660  
TCACTTGAAG TGAGTTTATC TACATTAAAG AAAATATCTG AAAAAACCAAT ACGAAGTTTC 720  
35 ATGTTTGATT CTAAATAATC GAAGATAGAG CCCTTAGCAA TATCATCATT TAAATATTTC 780  
ACGATTTCTT TATGATAATA AGAATATTCG ATACATAAAA CATCATCGTC CACGAATCTT 840  
AATCGCTCTA AATAGTAGAC GGTATCACTC GTGCC 875

40 (2) INFORMATION FOR SEQ ID NO:126:

(i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 654 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

55 CAACTAGTGA TTCAGGATAT AGGAATTGTA TTGTCGGAAT TCANACAGGT GGATATGGTC 60

EP 0 841 394 A2

	AAGATTTAAA	AAGATTATAA	CTTGGCCCAA	TTATNACGGN	GATCTTGAAA	ACGATTANTG	120
	GATTAGAACG	AATTCGAATT	TCTTCAATTG	AAGCAAGTCA	ACTTACAGAT	GAAGTAATTG	180
5	ACGTTT TAGA	ACGTTCAACA	AAAGTTGTGC	GTCATTTGCA	TATTCCATTA	CAATCTGGTT	240
	CAGATACAGT	ATTAAAACGT	ATGAGACGTA	AGTATACAAT	GGATAGATTT	TCAGAACGAT	300
	TAACAAAATT	GCATAAAGCT	TTACCAGACT	TGGCAGTTAC	GAGTGATGTA	ATTGTTGGTT	360
	TCCCAGGTGA	AACTGAAGCA	GAGTTCCAAG	AAACATATGA	TTTATCGTA	AAGCATAAGT	420
	TCTCTGAACT	GCATGTTTTC	CCTTATTCTC	CTAGAATTGG	CACGCCAGCT	GCAAGAATGG	480
10	ATGACCAAAT	TGATGAAGAA	ATTAAAAATG	AACGCGTGCA	TAAGGTAAAT	TACGCTTAAG	540
	CAATCAACTC	GGAAAGTTAT	ATGCGTCTTA	AATTGGATCA	AAGATGTGCT	TGGANGGTNA	600
	TTCCCTGAGG	GAACAAGGGT	GGACACAAGA	AGGGTACAAT	TAGGTGGRD	CTNS	654

15 (2) INFORMATION FOR SEQ ID NO:127:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 1443 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

30	CTANCNAANG	GAANTTCAGC	ATCCTTAAAA	ATACCTATTT	GACTGTAGAA	ACCTTTTGNT	60
	GCGTACAATA	TCTAAACCTT	GTCGTGCTGC	TGGAAC TGCA	CCTGAACATT	CAACAACAAC	120
	ATCTGCACCG	TAACCGCTCG	TAATTCCATT	GATATACGTT	TTTAAGCTCG	TGTGTTGTAA	180
35	ATTGACTACA	TAATCCATGT	GCAATGCTTC	TGCTTTATCT	AATCTGACTT	NGTGGCANTG	240
	TCCAATCCAG	TTACCACAAC	AGGTGCGCCT	TTACTTTTCA	ACACTTGTCG	TACAAGTAAT	300
	CCGATTGGCC	CAGGTCCCAT	TACAAC TGCT	ACATCGCCAG	AGTTCACTTG	AATCTTAGAA	360
	ACGCCATGAT	GTGCACATGC	TAATGGTTCT	TGTCATAGCT	GCAGACTGAT	ACGATACTTC	420
40	CGCTTCTGGA	ATATGATNCA	AACTTCTTTC	ACGTGCAATG	ACATAATTAG	TAAATGCGCC	480
	ATCAACTTGT	GTTCCAATAC	CTTTTCGATG	GTTGCATAAA	TGATAGTTTT	TTGATTTACA	540
	GGAATCACAC	TCATTACANA	CCATAGAATG	TAGTTTCAGA	AGTGACNCGG	TCACCAACTT	600
	TAAATCNTT	AACGCTCGCT	CCCAACTTCA	ACGATNTCAC	CAGAAAATTC	ATGACCTAAT	660
45	GTCACTGGAA	AATTAAC TTN	ATAATGCCCT	TCATAAGTAT	GAAGGTCTGT	GCCACAAATT	720
	CCTGCATAAT	GTACTTTAAT	CTTTACTTTA	TCATCTAGCG	GTGTTGCAAC	TTCTTTATCA	780
	AGAAAGTTCTA	AGTTGCCATG	TCCTTCTCTT	GTTTTTACTA	AAGCTTCCAC	CACAAACACN	840
	TCGANTTTTT	ANTTGNAATA	GACTNNATAG	NTTNAAGATA	AGATAGTTAN	CGATATTNCC	900
50	ACCTTGATCA	ATACTTGANA	TTTCAGATGA	ACCTTTTGNC	ATTGTACAT	TCGTACCTTT	960
	CGCCATATCT	GTGAAAATGG	GTGCTACGTC	TGTTGCAATA	TATAATGAAA	TTGCAATCAT	1020
	AATCGTACCC	ACAATGACAG	AATGAATAAT	GTTTCCTCTT	GCTGCACCAA	CAATAAACGC	1080
	GACAACAAAT	GGTATAGTTG	CTAAGTCACC	AAAAGGTAGT	ACTTGGTTTC	CTGGTAAAAT	1140
55	AACGGCTAAT	AAAACAGTGA	TAGGTACTAA	AATTAATGCT	GTCGAAATAA	CCGCTGGATG	1200
	ACCTAATGCT	ACAGCCGCAT	CCAATCCAAT	ATAAATTTCA	CGTTCGCCAA	AACGTTTATT	1260

# EP 0 841 394 A2

TAGCCATGTT CTTGCAGACT CTGAAACTGG CATTAAACCT TCCATTAAGA TTTTACCAT 1320  
TCTAGGCATT AAGACCATTA CTGCAGCCAT TGACATTCCT AAATTAATGA TGTCTCCAGG 1380  
TTTGTAACCT GCTAACACAC CAATACCTAA ACCTAAAATT AAGCCGACAA ATATAGACTC 1440  
TCC 1443

## (2) INFORMATION FOR SEQ ID NO:128:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1103 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

CTTAGCAACT CAAGAAAACG TTAATTTATC AATTCGCGGA ATTCCTAAAG AAGTAGTAGC 60  
ACGTGGTGCA CATCCATGTA TTATTTCTAT GGAGGGTCTT GAAAAAGAAG GCGACACTTA 120  
TGTCATTCCCT CATGTACATG AATTATTGAC ACCATTAGTT TCAGTAGTTG CATTACAATT 180  
GATTTTCATAT TATGCAGCAT TACACAGAGA TTTAGATGTG GATAAACCAC GTAACCTTGC 240  
TAAATCAGTT ACTGTAGAAT AATTCACCTT TTTAGAATCA ATCATGTATT AAAATTAAAG 300  
TATATGGCAC CCTTTTAGAT TAATCGACTA GAAGGGTGCT TTTTAGGTC GACTTAGCTT 360  
TTACTTTCATC TTAATTTGGC AGAAATGCGT NAAAAGGAAG TGTTTTATTT ATTTAAATAG 420  
TCTGACAATT AAGGGTGTTA TGTTAATATG ATTTTATGAG AAGTATGGAG TAGCNATAGA 480  
GGGGTGACCT CGCATGTAA TTCAATTAAG ATCCACATTG GGCGANTGAA GCAAGGAAAA 540  
ACAATTTTAA AAAATATTTT TTGCCACATT CCTAAAGGTG ATAAATGGAT ATTATATGGG 600  
TTGANTGGTG CTGGTAAGAC AACGCTATTA AATATATTAA ATGCATATGA GCCTGCGACA 660  
ACTGGCGGTG TTAATCTATT TGGTAAATG CCAGGCAAGG TAGGGTATTC TGCAGAGACT 720  
GTNCGACAAC ATATAGGTTT TGTATCTCAT AGTTTACTGG AAAAGTNTCC AAGAGGGTGA 780  
AAGAGTAATC GATGTGGTGA TAAGCGGTGC TTTTAAATCA ACTGGTGTTC ATCAAGATAT 840  
TGATGATGAG GTACGTAATG AAGCACATCA TTTACTTAAA TTAGTTGGAA TATCTGCTAA 900  
AGCGCAACAA TATATTGGTT ATTTATCTAC AGGTGAAAAA CAACGAGTGA TGATTGCACG 960  
AGCTTTAATG GGGCAACCTC AGGTTTTTAA TTTTAGATGA GCCAGCAGCT GGTTTAGACT 1020  
TTATTGCTCG TGAGTCACTA TTGAATATAC TTGACTCATT GTCAGATTCA TATCCAACGC 1080  
CTTGCATGA TTATGTGACG CAC 1103

## (2) INFORMATION FOR SEQ ID NO:129:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1095 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

```

CTGACATATT TAATGTTTAC ATCTAGTGCA TCAGGTTACG CCGTNTGTAT AGCGATGTTG      60
AGACGTTTAA CTGGATTAACTAATATCATT GGTAATTTT ANCCAAAANA TATTGNTNGG      120
GTTTAATTGG CCCGAGTACT TTTACCANTA TCATGGTTAA NTAGTATTTT GNTGTGACTC      180
NAGGTGTACC NCCCAAACCG TTGCATGCTA ATTTAATGAT TCGGACTTTA AGCGGACATA      240
TCCAACATAT TGCATTTGGA CCTATTGCAT CACTTGAATC AATAAAACAT CTTGGTACGA      300
ATGGTGGAGG ATTTTTAGCA GGAAATTCTG CAACACCTTT TGA AAAATCCA AATATTTGGA      360
GCAATTTTAT AGAAATGGGC AGTATGATGT TACTTCCTAT GTCAATGTTG TTCTTATTTG      420
GTCGCATGTT AAGTAGACAT GGTAAACGAG TACATCGTCA TGC GTTGATA TTATTTGTCG      480
CAATGTTTTT CATTTTATA GCAATCTTCA CATTAACTAT GTGGAGTGAG TATCGTGGTA      540
ATCCAATACT AGCGAATTTA GGCATTTATG GACCGAATAT GGAAGGTAA GAGGTACGGT      600
TTGGAGCAGG TTTGTCAGCA CTATTTACAG TTATTACGAC GGCATTTACA ACGGGTTCTG      660
TTAATAACAT GCATGATAGC TTAACGCCTA TAGGTGGATT AGGACCAATG GTATTAATGA      720
TGCTAAATGT TGTATTTGGT GGC GAAGGCG TAGGACTCAT GAATTTATTG ATATTNNTCT      780
TACTGACGGT GTTTATATGC AGTTTGATGG TTGGTAAAC ACCAGAATAT TTAAATATGC      840
CAATTGGCGC CCGTGAAATG AAATGTATTG TCTTAGTCTT TCTCATAAC CCAATTTTAA      900
TTTTAGTATT TTCAGCACTT GCTTTTATGA TTCCTGGAGC AAGTGAAAGT ATAACGAATC      960
CGTCTTTTCA TGGTATTTCA CAAGTTATGT TATGAAATGA CATCAGCTGC TGGCGAACAA     1020
TGGATCAGGG TTTGAAGGAC TGAAANGAAT GAATNCAACA TTTCTGGGAA ATATCTCTAC     1080
AGGAATCATT AATGT                                     1095

```

(2) INFORMATION FOR SEQ ID NO:130:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 913 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

```

GGTGTGTGTTT TCCTCCAATT TTTGGATCTT TTTTGCTCTA AAATCTCATC TTTTTTAATC      60
TAAATAGATC AAGAGTGCGC TCTAAGTTTT CTCTTTTGTG AAGAAATCGG GTTTCCTTGT      120
CTNTGAAGAT CNCAACGTCT TTCTCGNAGT TCTGCTTCAG TNTGTTCTCT TNGGAGTTGG      180
TNTTCTTCTT TTGCNTCAAG TAATTTCTCT TTTTGTGATN TGTCAGCTTC TTTATGTGCT      240
TGATNAACAA TATCNTCGGC AGTTTGCTTA GCTTGTGATT GCTTTTGAAG CANCANATTT      300
CGGCGNACAA CATACCTTCC AACAACTCCT AGAATGATCC CCANCACATG AGTAGGAGGC      360
TTAATAAATT CACACACACA CCTCCTTTTC TAGGGGTTTG CTCCTTTAAT ATGAAATCCA      420

```



# EP 0 841 394 A2

5 ATATGACTTT ATGAAATTAA AAATAAATCA AACCAAATTA ATTTGTGCGG TTTAATAGGA 480  
 GATAAGTGTC AAGTTATCAA TTTGCACTTT TGCACATAT TGAATGCTAT TTCTACTATT 540  
 TTCTAAAGTT TTGAAGCTCG ATTATAAAAC TATTGATAAT GTAATAATAA ATAAAATAAA 600  
 AAACCTTATA GTCTCGATTT GTAGTGTATC CCATAAAGTT AGATATTATT CTAGCTTTAT 660  
 GAGATGAATA TCACTTTCAA TCAATCCTAT AAGGATATTT GCTAAGTTTA ACTATAGATA 720  
 TAAATTTGTG TACTATTCTT CGTCAAATAA TGACTTTGGT GCATCTTCTG TTTCTTCAAC 780  
 ATCACCATCA GATATACCTA ATTTTCTCTT CAATTTACGA TCAATTTCTT CTTTAATTG 840  
 10 TGGATTTTCT TTCAAGTACA TTTTAACATT TTCCTTACCT NGACCCATTC GTTCGCCATN 900  
 GCAAGAATAC CAG 913

## (2) INFORMATION FOR SEQ ID NO:131:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 970 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

30 TGANGTAAAT TTGGTTGTGG CTCANTCGCA TCTGCGAAAG NGNAGTGCCC NGAGACCGCC 60  
 TGTAATGTTT ATGACGACCC TAGAATAATA ATAACTAAA AGCGCCGCCT NATGAAANTG 120  
 AGCGCCTTGG ATGNAATCAC TCCAAACGGA CACCNTCGAA NCCACCTNGA GATGTANANG 180  
 AAATACATAG TAAACCTACG AGTGATCCAN CGATATAAGG GTTCATGTCT GATACAGATG 240  
 35 TGATTGCTAA TGTTGGTAAG TAGATAACAA TTGCAACACG CCTTAAATGG TAAACGACAA 300  
 ATAATAATGA GCCANTGACA CGTATGCTAG GGCCAAATCT AGCTTCTAAA TATTCATATG 360  
 CAGATGTTAC CTTTAACTTT TTAAAGAAAAG GGACATAGAA ATAAATAAGT AATGGAATAA 420  
 TTGCGACGAT AGCAATGTTA CCAGCGATAT ATGACCAATC TGTTAAAAAT GCTTTCCTCT 480  
 40 GGTGTAGACA TAAATGTAAT CGCACTTAAC GTAGTAGCAT AAATTGAAAA GCCAACTACC 540  
 CAAGATGGCA AGCGACCACT TGCGGTAAAG AAATATTGG TACTTTGGCT CGCGCGCTTG 600  
 GTAAAATAAA CGCCAATGAA CAACATAGCT AGTAGATAAA TGATAACGGC AACCAGTTT 660  
 AGTGTGCCAA ATCCAACCTC TTTCATGGGC AACATCCCTT TTACAATGTA TTGATTCTTT 720  
 45 GATGTCTATA AATCGTATTT TGCAATGAGT TGATCTACAT GTTCGTCGAT GTGCTTCGTT 780  
 AAAAGGTTTG AAAGGTCTTT TCGGTAATCC TGCATCAATG TCACGATGAC GTAATATTTT 840  
 TTTCAATGTT GGATAAATCC CCATNGATNA CAACTGTTTC GATAATGTCG TTTGAATCAT 900  
 GTTGTAGTTG GTAAGCTTCT TGAATTTGAC CTTGTCGTGC TAAGTCGAAG ANTTCTCTAG 960  
 50 CGCTCGTGCC 970

## (2) INFORMATION FOR SEQ ID NO:132:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 791 base pairs

EP 0 841 394 A2

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

GAAGACATGA	ACGCATACGA	AGATAGTTAT	GAAAACCGAA	CTAAAAAAC	AACTGCTGGC	60
AGTAGCGATT	TAAGTATGTA	ATTAACGAGG	GCATTAGCCC	TCTATTTTGT	TGTACAAAAT	120
AACGATAAAC	GAGGTATTTA	ATATGACTGA	AACAACCTTT	AATCCAATTA	CATCATTAAC	180
AATTAACAAT	GAAGAAGTGA	AAGCAAAAGC	AACATTTATG	TTCGATAAAA	CCGCTAAAAA	240
ATTTGCAACT	GAACAAGAAG	ATAACAAAGG	TAGAAAACAA	ACGATTTTCAG	GATTTACTAA	300
TGTTTATNAC	GCTTTATTAG	AACGTGACAC	AGTGGCAATC	GTAGACTTTT	GGGAATGTGC	360
AACAGCTTAT	CTAGGTAAAA	GCGCGCCTAA	AAGAGAAGAT	ATTGAAGCGG	AAATCATGGA	420
AATCATCGAA	AGAGAAAAACG	ACACGTTGAT	CTTTTACAAG	GTGCGTTGGA	CGTAATGAAT	480
AATAGCGGTT	TTTTCAAGCA	GAAATCACGT	CTATTCTGGA	CACAGATGAA	CCAAGCGCCA	540
TCGTTAGCCA	AAGAAGACGA	GAAAGAGGGC	GCGAAAGCTG	GTATCGAGAT	GATGAAGAAC	600
AACTACAAAG	AAATCATGAC	CGTAGCACCT	TATTAGACTA	TTCGGAAATA	AGGCAGATGA	660
CAAGTCGTTA	CATAGGTTAT	ATGAGTAATG	ACGAGCTAAT	GAGCATGCTA	CCTGCCGAAT	720
GGAATGACTG	GATTATTGGC	GCTAGACAAG	CATTGATTGA	CCAAAGGGAC	NTCGCTCGTC	780
CCGCTCGTGC	C					791

(2) INFORMATION FOR SEQ ID NO:133:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 820 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

GGCACGAGCT	AAAATATATG	AAGCTAAAGG	CCGTCCATCT	GACAAATCCGC	TTATTGTTCA	60
TATACACAGT	AAAGTTCAAT	TAAAAGATTT	TACATATACT	TTGGCACGAG	CGTAGAAAAG	120
TTAATGCAGG	CATTCTGGCC	GGGCCCTATT	TCGTTTATAT	TGCCGCTAAA	GCCAGGCTAT	180
CTATGTCGAA	AAGTTTCTGG	AGGTTTATCA	TCAGTTGCTG	TTAGAATGCC	AAGCCATTCT	240
GTAGGTAGAC	AATTATTACA	AATCATAAAT	GAACCTCTAG	CTGCTCCAAG	TGCTAATTTA	300
AGTGGTAGAC	CTTCACCAAC	AACTTTCAAT	CATGTATATC	AAGATTGGA	TGCCCCGTATA	360
GATGGTATTG	TTCAAGCTGA	ACAAAGTGAA	GAAGGATTAG	AAAGTACGGT	TTTAGATTGC	420
ACATCTTTTC	CTTATAAAAT	TGCAAGACCT	GGTCTATATA	CAGCAGCAAT	GATTACAGAA	480
ATACTTCCGA	ATAGTATCGC	CCATGCTGAT	TATAATGATA	CTGAACAGGC	AATTGCACCA	540

# EP 0 841 394 A2

GGTATGAAGT ATTAAGCATT ACTCACCCAA TACAACACTT ACAATTATTA CAGATATTGA 600  
GAGCAAAATT GGAAATGACG GTAAAGATTG GTCTTCTATA GCTTTTATTG TGCCGAGTAA 660  
TAAGGTGGCG TTTATACCAA GTGAAGCGCA ATTCATTCAA TTATGTCAGG ATGACAATGA 720  
TGTAAACAA GCAAGTCATA ATCTTTATGA TGGTGTTACA TTCACTTGAT GAAAATGAAA 780  
ATATTTCAGC GCGGTATATA TACGGCTTTG AGCTCGTGCC 820

## (2) INFORMATION FOR SEQ ID NO:134:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1107 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:

GTGACTTGAA AGATATAGTA TATCGGATTT GTAATGCGCC GTTAATCAAT TATGATGCGA 60  
TTAAGATGAC CCGACTTAAA ATACNCAGTC CGACACAAGA TTGCTATGAG TTTTTTGGGT 120  
GGAGATATTG AATTATATAA TTATCACTAT TAATACCTAC ATCAAAGGTA TGCTATTGAT 180  
TTAGTTATTA AAGNAAATCA TAAACATAT AAGAATCAAG GAAAAGTAAA TACTGATTAT 240  
TTTTGTTTTG GAAAAGATAT CATTGCGCCA GCAAATGGTA CAGTTGAAAA AGTAGTGAAT 300  
GGGGTTCAAG ACAACAGTAT TGGAAGTACG AATGAATCAC AGTTTTTAGG AAATTATATT 360  
GTAATTAAGC ACGCAGAAAA TGAGTATAGC TTAATAGCTC ATTTACAACA ATATTCAATC 420  
ATTGTGAATG AGGGGCAAAA TGTTAAATAT GGTGACTTCC TTGGGAAGGT TGGGAATTCT 480  
GGCAATTCTA CTGAACCTCA TATACATTTC CAAGTAATGA ATGATAAGAA TATTGAAGCA 540  
TGTACGTCTT TGAAAATTCG ATTTTAAAT AATCTAGAAC TTATCAAAGG GGATGTGGTC 600  
TGCGGATTAC AAGGTGAATG ATGGCGATAC TTATAAAATC TCGACACTAT AAAAATGGTA 660  
TAGTGTCTGAG ATTTTCTTGC TTATTTAGTT AATTCAAAGT GTACGCCGGA TTCATTAGAA 720  
GTTGATGTAT TTGTGTTTGT AATAGAGTAT CCAGTTATTG AAATTTTAGA TTCAATATCT 780  
GAAGCGGTAT TTGTAGATTC GGGATTATAG AAGCTACATT CATAAGTGTT ATCATCTNTG 840  
GTNTTCNAAG AATTAAACAT TCCCTTTGGC TTTTAACTTC CGACTTTAGT GGTGGGNTAA 900  
TGNCAAAAAG TNTGGAGNTG CTTATTACTA TTAATTAACA ACCCAGCCCA ATTGGGNCAC 960  
CCCCAACGG ATTTGGACCA ATCCTCAACT TACCTGCTTA TGAATCCCTG GCTCCCAACC 1020  
CCCGCAAATT TGCTAAACCC AAAATCCACT TCCAANGTTT CCAAACCTTAC TTTCTTCCAA 1080  
GNTAATTTCT TCAAGGATTT TTTTGG 1107

## (2) INFORMATION FOR SEQ ID NO:135:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1135 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

# EP 0 841 394 A2

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:

	AAGTTTGATC	CACCCACCNT	AAACCCATCG	CCTAACAATA	AANCCGAANN	TCACCTCCTG	60
10	CCTGTTGATC	CAATTGTTCC	CATTCCAATA	ACCGCCACCT	CTTTTCTAG	AACACCCACA	120
	CCTATCGTTC	TATGGTTAAT	ATGTTGACA	TCTGTATGCG	GTATAGCGAC	CGAACATAGA	180
	TGTGTTGGTA	AACCAGTANC	NAATTCTTTT	TCTCTGTCGA	TGACTGCATC	TTTAAACGTT	240
	GACTTCACGA	ACCCATTTTG	AAATAACACA	TCTGACATTT	GNGACAATAC	GGATTCTTTA	300
15	TCAGTTGCCG	ACNAATTGAG	CATTATATTT	TCTTTATGCA	CCTAATTGCT	TGTCCCATCC	360
	ATTTTCCCTC	GCTTCTTTAT	TTGAATAATT	TTTTTAAATC	TCATTTACAT	CAGAATTTTT	420
	GCGACTTGT	ATGATGCGCT	TAATTGCGTC	ATTGTCCTGC	GCCACATCTC	TCAATTGTAG	480
	TAACGCTCTT	AAGTGTGTCA	CTTTATCAAC	AGCAGCAATA	GGTACAATAA	TATGGATTGC	540
20	TGTGCCATCT	GACATGTATA	TTGGTTCTTG	TAATATCAAC	ATACTCATCG	CTGTTTTATG	600
	TACATGCTTT	TCAGAGTATG	CATGTGGAAT	AGCAATATTT	TGCATGATAA	CCATATACGA	660
	ATCANCANAT	ACCNGGGAAT	TCCATCTCAT	GAATATATTT	AATATCAATA	AAATGATTAG	720
	CAACTAACAC	AGCACGNGCT	TTAGCAATAG	CTTCATCAAT	ATTTTCAACA	GGAGGCNTNC	780
25	TTTNCACGGG	CCTGGCCGGT	ATCAAGGCAG	CTAAATCTAA	TGCTTTATTT	TGTGTGACAA	840
	TCGATCCATT	AATGGTTGAA	ATTGAATTAT	AATTGGCAAT	AAAATCTTCT	AAACCATCAC	900
	GTAGTCTGT	AATGTCATTA	ACTGTCGTTG	TGCGTTCAAT	TAATGCCATT	AACCTGTTTA	960
	TTTCCTTATC	AATGTCAGCC	GATTCCTTAT	TAATGTACTT	CATCACTTCT	TTACGTAAC	1020
30	TTCGTTGCTC	ATTTTCAGAT	AAAGTFACTT	TTGTGATAAA	TAATTTTTTA	TGTGTTAGGA	1080
	CAACATTTGG	TGAAAAGACG	AAGTCATAAT	CTAATGTGTA	ATTTTCAAAT	GTTCT	1135

(2) INFORMATION FOR SEQ ID NO:136:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 973 base pairs

(B) TYPE: nucleic acid

40

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:

	CGTGCGAGCT	GCTTTTGGTA	AAAACGGCGG	TAATATGGGT	GTATCTGGAT	CAGTTGCTTA	60
50	TATGTTTGAT	CATGTGGCAA	CATTGGGTAT	TGAAGGAAAG	TCTGTTGACG	AAATACTTGA	120
	AACATTAATG	GAACCAAGAT	GTAAGATGNA	AATGATGTGA	TTGACGATAA	TGGATTGACA	180
	ATAGTCTATG	CTTGAACCAA	GATCCAATTT	CCAGTCGTTC	CAAGATGCGC	TTCGTNCACC	240
	AGGTGTTGG	AAGAATTTAA	AAGTTGCTTG	AATTTGAAAT	TGTTACCCCTC	CAAACAAGAT	300
55	ATTGAACTTT	CTGAAGCGGA	CCNAAGTAAC	ATTTGAAAAA	TTAATCGATG	CATTAGAAGA	360

# EP 0 841 394 A2

	TTTAGAAGAT	GTACAAAACG	TATTCCATAA	TGTGGATTTG	AAATAATGAA	ATCAGCAGAA	420
	CAATGGATTG	ATGAATTGCA	ACTTGAATCA	CATCCTGAAG	GTGGTTTCTA	TAGAGAGACA	480
	ATTCGAGAAG	TATTGAAAGA	TGGACGCAGA	GCGCCGTTTA	GNAGTATTTA	TTTTTTACTT	540
5	ACAGATGACA	ATATTTGCGA	TTTTCATCGA	ATTGATGCTG	ATGAAGTATG	GTAATATCAT	600
	GCTGGTGATT	CTCTAACAAT	TCATATGATA	AATCCGGATG	GGGAATATAC	GAATGCAACA	660
	TTGGGTACTG	ATATCCAAAA	TGGAGATGTA	TTGCAATATG	TAGTGCCTAA	AGGAACAATT	720
	TTTGCTTCTT	CAATCGAAAT	TTCAAATACT	TATAGTTTAG	TAGGTTGTAT	GTGTCAACCG	780
10	GCATNTGAGT	TTAAGCAGTT	TGAATTGTTT	AAGCAATCTG	AATTAATTAC	ACAATATCCG	840
	CATCTTAAAT	CAGTGATTGA	NAAATATGCT	TTAAAAATAA	AATGATCAAT	GAAGTGGTTT	900
	GAAGGTTGTT	AATAAACCTT	TGAGTCACTT	CATTTTATA	TGTATTCTTG	ATTGAATCAG	960
	AATAGATTTG	ATG					973

## (2) INFORMATION FOR SEQ ID NO:137:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1701 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

	GGCACGAGCT	GGTTGCATAT	CTAAACTTTG	GTAATCTACA	CCTTCTTTAG	ATTTACCGAA	60
	TGCGATAGCT	GGAGACTTAA	CTGCGGCTAC	TTTCGCATGT	GGAATAGCAA	TACCTCGCCG	120
	ATACCAGTTG	TACTTTGTGA	TTCTCGATTG	TGAATCGCTT	CCTTAAATGA	CGCGACATCA	180
35	CTTAATTTAC	CTGCCTTGNC	TAATTGGAAT	TACCTAACTC	ATCAATAACA	CCATTTTGTG	240
	CATTTGCCAN	TAAATCCATT	GCTATCGTAT	CTTTTGTAA	TAACTCTGTT	ACTCTCATT	300
	TTTTCACCTC	CCATCAAGTA	CGCTAATCGN	AACTTGGTGA	TTCTATTTT	TCTTATAGCG	360
	GTCCCGTGTT	TGCTAAGTCC	TCAATCAAAA	TGCCGGGGCA	AGTACCGCAA	TGCGACTGCT	420
40	TGGTTGGGAA	TGCTTTTTC	ATCGTTAAAC	CTGGAAGCAA	TTCCAGCCAC	CATGCCTGCA	480
	ACTGTACTAT	CACCAGAGCC	AACTGTATTA	ACCACTTTCC	CTTGTGGATT	AACTGCTTTA	540
	ATACTGATTT	CTTTATCAAT	ATAAATAGCA	CCATCACCGC	CAAGCGAGAC	AATAACAGAT	600
	TGCGCACCTT	TATCAACTAA	CAAACGACCA	TATTTAATAA	CATCTGTGTC	TGAGTTCACT	660
45	GTTGTATTAA	ACATCACTTC	TAATTCATCT	TTATTAGGTT	TAATAAATAG	TGGNTGAAAT	720
	GGTAAAACGC	TTTCAAGCCA	ATTCTTTTTC	AGCGTCGACC	GACTAATTTA	GCACCTGTCT	780
	GTGCTGTAAT	TTGTGCAAGT	TTGCGCATAC	GCATCGCTTG	GAATACNACT	TGGTACACTT	840
50	CCAGCAACAA	TAACGATATC	NTCGCTTGTT	GTATTTTAA	TGTGTTGTAA	CAGTTGTTCA	900
	AATGTGTTG	ACGTTATATG	AGGACCCGGT	GCATTGATTT	CTGTTTCTTG	TCCTGTTTTT	960
	AATTCACAT	TAATACGTGT	ATCTTCATCN	ACNNCAATAA	AATTCGATTG	AATTGCNCTG	1020
	TTANTTAATG	TATCTATAAT	GAATTTCCCA	GGAAATCCAC	CTGCANATCC	CAAGGCAGTT	1080
55	GACTCAACAT	CTAATGTCTT	TAAGACGGCG	CGAGACGATT	AATACCCCTT	CCCCCCCAGC	1140
	GAAGTAATAT	GTTGCTGTTG	CTCTGTCCAA	NNCCATCAAG	GTTNAGAATC	ATTCGTAAAA	1200

# EP 0 841 394 A2

ATGACATAAG TCAATTGTAG GAGTGNGAGT CNCCTGTATA AATCATAAAG TCCCTCCTAT 1260  
 AAAGTGAGAC TTTTGTGGT ATTCTTTTAN CGAGTCTTGA GTTAATGCTT TTTCAGATGT 1320  
 5 GATGANTGTC GTACTTTCTA GCACAGGGAC ACGAGCACAA TATACTTTAT TAAACTTAGA 1380  
 NTGATCCTAT AAGNACANAT GANTGAGTGG CTAATGACAT TGCTGTTTGT GTAACATAATN 1440  
 CCTCTTGCTC ATCGGGAGTA GTTAATCCAA GTTCAATATC TAATCCATTG ATCCCGATAA 1500  
 AAGCTTTATC GNAACAATAT CGTCTTAATA TCTCCATAGC ACTAGANCCA ATCGTAGCNA 1560  
 10 GTGTATTTTC TTTAACTTGA CCACCTAGCA TAATTGTTT AATACCTTTT GGAAGTAAAG 1620  
 CTCTACATG TGTAAACCA TTGGTTACCA CAATGATATC TTTCGCTTGA ATATATTAA 1680  
 TTAGCGGCAC GAGCTCGTCC C 1701

## (2) INFORMATION FOR SEQ ID NO:138:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 858 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:

CCCGGTGTAA ACCGTATGGA TGGATGCTAG CACCAANTTA ATAATCCTCC ACAATANCAT 60  
 30 TAAGGGTATT CATGTAGGTG NCACAACATG GTAAAGGCTC TACCAGTTGC TTACCTTAGA 120  
 ACAGCTTTAA GTTGNAATG GTTATGAAGT AGGTACATTT ACGTCGCCGT TTATTGAAAC 180  
 ATTTAATGAN CGAATTAGTC TAAATGGTGT GCCAATATCA AATGACGCTA TTGTAGAATT 240  
 35 AGTATCACGT ATTAAACCAG TAAGTGAAAT GATGGAACGT GAAACAGATT TAGGTGTTGC 300  
 AACTGAATTC GAAATAATCA CAGCGATGAT GTTTTATAT TTTGGTGAAA TACATCCTGT 360  
 TGATTTTGTC ATTGTTGAGG CTGGATTGGG TATAAAGAAC GATTGACAA ATGTCTTTAC 420  
 ACCGGTTTTA TCAATCTTAA CTAGTATCGG TCTAGACCAT ACAGATATTT TAGGTGGTAC 480  
 40 TTATCTAGAT ATTGCTAGGG ATAAAGGCGC GATTATAAAG CCTAACGTTT CAGTGATATA 540  
 TGCTGTAAA AATGAAGATG CATTAAAATA TGTTGCTGA ACGCGCAATT GAACAACATG 600  
 CAAAGCCAAT TGAATTAGAT AGAGAAATTG TTGTTGTATC GCAAAAATGA TGAATTTACT 660  
 TACCCGNTAT TAAAGATTAT GAATNAGAAA CAATCAATNT NAAGCATTGT TAGGGTNGAA 720  
 45 CATCAAGAAA CAANATGCTG CATTAGCCAT AACAAANTCTT ATTGGANTCA AATTGAACAA 780  
 GGATTAATTG AATNCAGATT TNCAANAAGA TGATAGACGG GTTTGATCAA GTCCGTGGAC 840  
 NTGNCGTTTT GGCAGGT 858

## (2) INFORMATION FOR SEQ ID NO:139:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 497 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

# EP 0 841 394 A2

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:139:

GAATTCNAAG	TCACCTTTCTT	TTTTAAGCTT	TAAATTTCTC	CCCATTTTTT	TAGCCCCCTA	60
TAAGGATTGA	ATATCAATGC	CTTCTNTCAT	TAAAATTTCT	CTAATTTGCG	AAACAAATAA	120
TAATGCATGT	TCTCCATCAC	CATGCACACA	AATTGTATCT	GCTTGTAACG	TTACTTCCTT	180
ATTGTTTTGT	GAAATAACTT	TATTTTCCNN	CACCATCTTT	AAAACCTGCT	TAAGTGCTTC	240
GTCAGTATCA	GTAATCACAG	CATCACTTTC	TTTTCTGATT	GATAAAATCA	TATATGTTCC	300
TATAAACACT	AAAAATCCTA	TAAC TAGGTA	ATAATATTAA	ATTCAAGGAT	CGANCCTCCG	360
CTAAGCGACA	ACAACAATGG	TGACAATAAC	AACAACAAC	GAGAATTNAA	ATACCTAATT	420
CAAAAAGGGG	TNATNGNCT	TACAATGAAT	GTGGCCNAAA	ATTGGGGAGG	NTCCAAGGNG	480
GGNTCAAATT	RRDCTNS					497

(2) INFORMATION FOR SEQ ID NO:140:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 969 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:140:

CCCTGGTTGA	TTGGAAAATG	TAATTGAATG	AAACGGGTAA	AACCTTCCCC	AAAGCACCAT	60
AATTGTTGCA	AGTAGGTTGC	CACCCCAATC	CTGGTTTACC	CTGATTTGAA	AGGCTTTGGG	120
CAAGCCGAAA	TATAGATATC	TCTAACCGAC	GAGATTTATA	ATTTGATGAT	TTATCAATTA	180
GGTGCATTAC	AAGGGTTTTG	TCGCATTCAT	CAACTTAAAA	TTAATCATGT	TAAACCCGCA	240
TGGTGCATTG	TATCAGATGG	GTGCAAAAGA	CAGAGAAATA	GCAAACCCCC	GTNATANCAC	300
NACCTGTTNA	TGNCNTTGAT	CCATCACTAG	TGTTAGTAGG	ATTAGCAATC	ATATCTAATT	360
TCAGAAGCAA	AGAATGTCGG	ATATAATCCA	GCTTCTGAAG	TGTTTGCTGA	TAGACGATAC	420
GAAGATGATG	GGCAGCTCGT	TAGTAGAAAA	GAAAGTGATG	CTGTGATTAC	TGATACTGAC	480
GAAGCACTTA	AGCAGGTTTT	AAAGATGGTG	NAGGAAAATA	AAGTTATTTC	AAAAACAAT	540
AAGGAAGTAA	CGTTACAAGC	AGATACAATT	TGTGTGCATG	GTGATGGAGA	ACATGCATTA	600
TTATTTGTTT	CGCAAATTAG	AGAAATTTTA	ATGAAAGAAG	GCATTGATAT	TCAATCCTTA	660
TAGGGGGCTA	AAAAAATGGG	GAGAAATTTA	AAGCTTAAAA	AAGAAAGTGA	CTTTGAAATC	720
ACAAAAAATC	ATAAAAGGTT	ATTATTAGGT	TCTGTATTTC	ACGATGGCAA	CTTCTGCAAT	780
TGGCCAGCA	TTTTTAACGC	AAACAGNAGT	ATCAACATCA	CAATCGTTTG	NAAGTCNCGN	840
ATCTGCCATA	TNACTGTCTA	TCATCATTGA	CATTGGTGCA	CAAATTAATA	TATGGCGCAT	900
ATNAGTTGTA	ACTGGTTTAA	GAGGTCAAGA	AATATCAAAT	AAAGTTGTTC	CTGGGCTTGG	960

... TCTCGTGCC

969

## (2) INFORMATION FOR SEQ ID NO:141:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 632 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

GGNACGAGCG	GCACGAGCGC	ACTTATTTTA	TCAGAACATT	CTAAAAAGAA	TATATTCAAT	60
ATAAAAAAGG	CTGTGTAGGG	GTGTCTTATG	AATAATAGAA	ATGTGTATGA	TATCGAAGTA	120
AGTGATTATA	AAGGCTTAAC	TTATAAATTA	GAAGCATTTA	GAGGTAAAGT	GATTTTAGTT	180
GTTAATACTG	CAACAGAATG	TATATATAGC	GAACAATTGA	AAAACTAGA	GACTTTGTNT	240
CAAAAATATA	AGGATCGTGG	GTTTGTAGTG	TTGAGTTCCC	CGAATAATAA	TTGTGACAAT	300
CGACAACCAG	GNTCTAATGA	AGAAATCTTG	AAAATTAGNC	GGNGAGAAAT	TTGGGTGTTA	360
CATTTCCAGT	GCTAGCTTAA	AATATCTTGT	GAACGGGAAC	GAATGAACAT	CCCCGNTATT	420
TACGCATTTN	ANAGGAATGA	ACAAACCAGG	AATCTTNGGG	CCCCCAATAA	AATGGAATTT	480
CACAAAAATT	TATAATCGAT	CGACAAGGCA	ATGTTGTAAA	TCGATTCTTG	CCAATGGAAG	540
ATCCAATGGA	TATATCGACA	AATATAGAAA	TATTATTGGA	AGAATCTTCA	ATCTTAAATT	600
TAANATTGAG	CGCTTAGTNT	GCAAATACAC	AA			632

## (2) INFORMATION FOR SEQ ID NO:142:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 662 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:142:

CCCCACNAAG	TACNCGAAAC	GCAAACAAAC	ATCTTAAAAG	GAGGAACGAA	CAATGCAAGC	60
ATTACAAACA	TTTAATTGGA	AAGAGCTACC	AGTAAGAACA	GTGGAAATTG	AAAACGAACC	120
TTATTTTGTA	GGAAAAGATA	TTGCTGAAAT	TTTAGGATAT	GCAAGGGCAG	ACAATGCCAT	180
CAGAAATCAT	GTTGATAGCG	AGGACAAGCT	GACGCACCAA	TTTAGCGACA	GCAGGTCAAA	240
ACAGAAATGT	AACGATCAAT	CAACGAATCA	GGATTATACA	GTTTAATCTT	TTCTAGCAAA	300
TTAGAAAATG	CGAAGCGGTT	CAAACGTTGG	GTAACCTCGG	AAGTTTACC	AACATTAAGA	360



EP 0 841 394 A2

AAAACTGGTG CTTACCAAGT ACCTAGCGAC CCAATGCAAG CATTGAGATT AATGTTTGAA 420  
GCTACAGAAG AAACAAAACA AGAAATTAAG AACGTAAAAG ATGATGTTGA TNGATTTGAA 480  
AGAAAATCAA AAACCTGGATG CGGGAGACTA CAATTTCTTA ACTAGAACTA TCAATCAAAG 540  
5 AGTAGCTCAC ATNCAAAGAC TACATGCGAT AACAAATCAA AAACAACGTA GCGAATTATT 600  
CNGGGATATT AATTCAGAAG TGAAAAAGAT GACTGGTGCA AGTTCAAGAA CGAATGTTAG 660  
AC 662

(2) INFORMATION FOR SEQ ID NO:143:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 960 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

TAATTCGCAA TAGGAGTGAT GAATATCATA AATTTTACCC TCCAAATGAA GCTAATGAAG 60  
TCCTGGACCC GAGTAAGACG CATGTAGCCA AGCTAAAATA ATCCACTCTA CCTTATCTTT 120  
AGTTAATAAT GTTACTAAAT GTTGTCATA CGCTGCTTTT GAATCAAATT GTTTTGGTTC 180  
ATTAATATAA ACAGGAATAT CGTGCTTGTT TGCTCTATCT ATACAAAACG CATTTTGATG 240  
30 ATCCGTATAT AGCNCCGTAA CTTCAATATT TTCAAGTTT CCTGATTCAA CATGCTCAAC 300  
TATATTTTCA AAGTTACTTC CTGAACCTGA TGCAAAAATC GCAATTTTAA CCATTGTTAT 360  
ACCCCAACA ATTCAATTGC AGTTGACTCA TTTTTCACAA TATGACCAAT TTGATAAGCT 420  
TCCACATTTT GTTCTGCTAA AATCTTCAAA GCGCGTCGAT GCATCTTTT CATCAACGAT 480  
35 AACCGTATAG CCAATACCCA TGTAAAAAT GTTATACATT TCATTGTGT CTATATTGCC 540  
TTGTTGTTGT AACCAATCAA ATATTTTGG CGTTGGAAAT GATGTAGTAT CAATTCTAGC 600  
AGCATATCCG GCTGGCAATG CACGTGGAAT ATTTTCATAA AACCTCCAC CAGTAATATG 660  
ATTCATTGCC TTAATAGAAA CTTCTTTTTT TAAAGCAAGT ACAGGTNTGA CATATAATTT 720  
40 AGTTGGCTCT AAAAAGACAT CTATAAATGG ACGATTATCG NAGGGTGATG CCAAATCAAT 780  
GNCTGATTCA NTAATTAATN TGCGCACTAA ACTGTNTCCA TTNGANTGAA TGNCACTTGG 840  
ACGCAAGTCC TATAACAAC TGGCCCTCTT NCAATCTTG AACCATCTTA CAATAGNCAA 900  
CCTTTTTCAA CTGCTCCAAC AGCAATCCG GCTACATCAT ATTCACCTTC GTGATACATT 960

(2) INFORMATION FOR SEQ ID NO:144:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1013 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

EP 0 841 394 A2

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

5  
 TAATNAATTT GCGGANTCAC TTGTCGCTAA GTGGCTCCAC CTTGTCATCT ATTATTGCAT 60  
 CACTCATGGG NCGNACATCA TAGTGTTGAT TTGATTACAG CATATCNACG TTTTGATTTT 120  
 10 TCTAATAGAA GATCAGCAAC AACATCAACA TTTGAATGAT TCATATATGA TGCAGGTACG 180  
 TCTTTTAATG TTTTAATGTT ATCAATATAA AGATTGATGT AGTGTTGCGG GATATTGTAG 240  
 TGATGTTCAA GTAACATATC AGTAACAAGT TGATTAAAGA CACTTTCATC TAATTCACCA 300  
 CGTGCCACAG CGCTTCTAT TAATGCTTTA TTTGGGAAAA TAGGCACGAG CGAACGTCAC 360  
 15 GTAACCATTT NCGGACATCT TCAAACGTAT CCGCTTCTAA TCCTTCCCAG GGGTTACGTG 420  
 CTGCAAAAAT CGAAATCGGT GATAATGGTG TAATAACACG TTTCGCATTT TCAATGACTG 480  
 AATTGATATT TAACTGTGTT GTCATACCTT TCACCTCCTA TAAATACTTC TTCAAATAAT 540  
 TCGGATGACT TTCTATCGCT TTCGAGCGTG CTTCACCTAG ATTAACCTAAC CACACGTACA 600  
 20 ATACCGCAAA ACCCTTAGAG TATCGATGAC GCGCCACCCA AATACTTAAT AAACGTCCAA 660  
 AGATTAAAAT GACAACACTA ATGATGACAC TCACTGTAGG CCGTGTTGTC GCATGTGTTG 720  
 TTATATTTTG GTAATACATC GTAAAAATAA TTGTGTGTGA TGACGTAGAT AAATGTCACA 780  
 ATTGCAATCA AAATCATGCC AACCGACGTT GCCATGCGCC CTTTACTAAA GGCTACCATC 840  
 25 TGATTCCAAG ATACAAGTTA ATGACCATGC TAGAATGAGT GCACTTAACA CTTTCATATGC 900  
 ACTTCTGTCA CTACCATCCA AATAGAATGC ACGATAATAG CTAATACACG TCCATGACAT 960  
 CCAGCATAAG CTCTTACAAT GCTGTTTTGA ATATGATCGC TCCCAARRDC TNS 1013

(2) INFORMATION FOR SEQ ID NO:145:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1032 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

45 ATCCACCCAT NACGNTACTC ATANCATTGT CAACAGTAAG TAATGTCCAT ATGCATATCN 60  
 ATNCCGCTCG GCANTATANC ACCCAACCGG ATTTTCATAAG TGGGCGGATA AAGTGATNAG 120  
 AGTGATCTCA AACTATCGA NTAGNTACCA GCGTTCCTTA TCGGTNTTAT TTTATNATTT 180  
 ATTGTCACAN ATTTANTGAA TATAGATAGC GTTATACTAA GTCAGNTTAT ATTACCTGTA 240  
 50 ATCACGCTAT CTTTAGGTAT GTGTGCATAC ATCATTCGTT TAGTGCGTTC TAATTTATTG 300  
 ATGTTATTGC AAAGTAATAT CGTACAANCA GCAAGATTAC GCGGTATGAA TGANC GTTAT 360  
 ATTTTAATTC ATGATTACT AAAACCAACA ATTTTGCCGA TTATCCCAT ACTAGGGATT 420  
 TCACTTGGCA GTCTAATAGG TGGTACTGTA GTGATTGAAA ATTTATTGTA TAATACCTGG 480  
 55 TATTGGTTAT CTATTAATGG GATAGTATTA AATCTCGAGA TTATCCTGTT ATTCCANGGAT 540  
 GCGTGTTATT TATTGGCTTC TTCGTTGTTA TTATCAATAC GATTGCTGAT TTATTAACGT 600

EP 0 841 394 A2

```

TATTACTTGA TCCGAAGCAG CGTTTACAAT TAGGAAATCC CACAAAACAC AACCAATAÇA 660
CCATTGATAT CAGAAAGTAG TGNCCGTCAT GCATAAAATA TTTTCANAGA ATAACCTGAT 720
ATTTTTTGTA TTCGTTGCAT TTATTTTGTG GGTAAATTGTA CTGCAATTCT TTGTCAGTAG 780
5 TGAAAATGCA ACCACAGTCA ATTTATCACA AACTTTTGAA CCGATTAGTT GGTGTCATTT 840
ATTAGGAACT GATGATTATG GGAGAGATTT ATTTACCCGA ATTATTATCG GTGCACGTTT 900
AACATTGTTT GTTACTGTTT TAACATTAAT AGCTATCGTT GTCATAGGTG TTACACTAGG 960
TCTATTTGCC GGATATAAAA AAGGGTGGAT TGAACGATTA GTGTTAAGGT TTATTGATGT 1020
10 TGGTCTAAGT AG 1032

```

(2) INFORMATION FOR SEQ ID NO:146:

15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 646 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- 20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

```

GGCACGAGAC TTTGGTGGTA GTTTTGAAAG ATTACAAGCA TTGACAACAA AAACAACATT 60
ACCCGTATTA TCCAAAGACT TTATAATAAG ACCCCCCCCA ANTTGATGTT CCTAAACCNA 120
30 GCTGGTGCAT CTATGATTTT ATTGATCGTT AACATCTTAT CGGATANACA ATTGAAAGAT 180
TNAATATACC CTACGCAATA TCCCAAATC NAGAAGTGTT AATTGAAGTA CATGATCCCC 240
ATGAATTAGA ACGTGCCNAT ANGGTTAATG CTAATTGAT TGGTGTAAT AACAGGGACT 300
TAAAACGATT TGTCACAAAT GTGGAACATN CAAATACTAT TTTAGAAAAT AAAAAACCAA 360
35 ATCATCATT TATTTCTGAA AGTGGTATTC ACGATGCATC TGATGTAAGA AAAATCTTGC 420
ATAGTGGTAT CGATGGCTTA CTAATAGGTG AGGCGCTTAT GCGTTGTGAC AATCTATCTG 480
AATTTTACG ACAACTGAAA ATNCNAAAGG TGAAGTCATG ATGAAATTGA AATTTTGTGG 540
CTTTACATCA ATAAAGGATG TTACAGCGGC CAGTCAATTA CCTATTGATT CGATAGGTTT 600
40 CATCCATTAT GAAAAAAGTA AAAGGCATCA AACAATACCC AAATAA 646

```

(2) INFORMATION FOR SEQ ID NO:147:

45 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 565 base pairs
- (B) TYPE: nucleic acid
- 50 (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

# EP 0 841 394 A2

GGCACGAGCG TCAGTCCAAA TCACGCCTTG TGGTTTCTCT GGAATTGTCA TCTATTACTC 60  
 5 ACCCCCAAAT ACATCATTA TGTAAATATT GTGAATTGCT TCAATTGGAT TTATTGTTTC 120  
 ATCTACAGNN CGATATCGCT TACTATCAAT CATGNCATCT ACATGACATA CCGATTGATA 180  
 ACTACAAAAA GCACATGGCA ATTTGTGTGT GTNCTTTAAT GGTGCAACTN CAGTATGTCC 240  
 ATCCATAATA TNTGAAGCTG TNTCTATAAA ATTCTCTNTG TTATGCTGAA TGAATTTATA 300  
 10 AATTGNTGGT TCATCTGACA CTTGGCTGGC TCGTTTACTC AAAGAGACAT CTTTATTATCA 360  
 ANCCAAGTGG GTACAAATAT CTGAAGGTGA ANTTAGGGTT CTTNAACGAA TTTCCAATGC 420  
 ATCAAATAGC AGGGGGGTCT TGCATTCAAC TAAACCCTCA GGTAAANCCT NNNAATTAAT 480  
 CTTGNTCTTG GNTNTCTTTC ANCAATATCA AGACCATGAT TTAATTTACT CTTGGGCCAA 540  
 15 GTCATGGAGT TTTATATCCA CCGGG 565

(2) INFORMATION FOR SEQ ID NO:148:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 919 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- 25 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

CTTGAAGTAG TAGGGTTACA AGGCTCAACT TACCTTTTAA AAGGACCAA CCGTGAAAAC 60  
 GTAAAGTTAA ACCAATCAGA AATGAACGAT GATGATGAAT TAGAAGTAGG TGAAGAATAT 120  
 35 AGTTTCTTCA TTTATCCAAA CCGTTCAGGT GAATTATTG GAACTCAAAA TATGCCTGAT 180  
 ATTACGAAAG ATAAATATGA TTTTGGTAAA GTACTTAAAA CGGATCGCGA TGGGGNACGT 240  
 ATAGATGTTG GGTTTACCCC GNGAAGTGTT AGTACCATGG GAAGATTTAC CAAAAGTGAA 300  
 ATCACTATGG CCACAACCTG GTGGATCATT TGNTAGTTAC ATTACGAATT GACCGTGAGA 360  
 40 ATCATATGTA TGGACGTTTA GCGAGTGAAT CTGTTGTAGA AAATATGTTT ACACCTGTAC 420  
 ACGATGATAA TTTAAAAAAC GAAGTCATTG AAGCCAAACC TTGGCNCGAG CGTATTACGA 480  
 ATTGGTAGCT TCTTATTAAG CGAATCAGGT TACAAAATTT TCGTACATGA ATCAGAACGT 540  
 AAAGCTGAAC CAAGATTAGG TGAATCTGTT CAAGTTAGAA TTATCGGGCA TAATGATAAA 600  
 45 GGTGAGTTAA ATGGTTCATT TTTACCACTT GCACATGAAC GNTTTAGACG ATGACGGCCA 660  
 AGTCATCTGT GATTACTAG GTGAATATGA TGGGGAATTA CCATTCTGGG ACAAATCAA 720  
 GCCCTGAAGC GATTAAAGAA GTATTCAAAT ATGAGTAAAG GTTCATTCAA AACGTGCAAA 780  
 TCGGNCACCT ATATTAAACA GAAGGATTAT TAATATTAGG AAACAGGGTA AAAATCACTT 840  
 50 TTAACTTAAA AAAGGGTTGG CGGGTCGNNT TTGGNCTCAA AAAGAATTAA TCATTNTNAC 900  
 AACGNNTTCG GNGGATGCG 919

(2) INFORMATION FOR SEQ ID NO:149:

55

(i) SEQUENCE CHARACTERISTICS:

# EP 0 841 394 A2

(A) LENGTH: 955 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:149:

CTGCATCTGG	TGAAGGAAAT	CAANTANNTN	TTATNNCTGA	AGTGATTGCA	TGGTGCAACN	60
TCAAATATCA	CATATGGCTC	AGTGGANTAT	ATGGATAAAG	GCTTAACAGG	TCATATCATG	120
CGGCGTGGTA	TTACTGAAGC	GGATGCCTCA	ATTAATTGGG	CACTAGGTTT	AATGAATGAG	180
GGTAGCCAAA	TTATTGATAA	TACAACAAAT	TTATGTGGTG	ATCGCACAAAC	AAGNNCACTT	240
AAATCAGGAG	GTGNAGGTAC	AGGAGAACAA	AAAATTAATC	TAACATCTAA	ANTCGCACAA	300
ATATGGTAAA	GAAACAGATG	GTTATNTCCT	TAAACATGGT	GTTATGAAAG	AACATGCATC	360
ATCTTGATTT	TTAATGGGTA	TCCGGCTACA	TTANGCATGG	GGGGAACATA	ATCAAGTGCT	420
AATCAGGAAT	CACGTGTTAT	TNATGTTATC	TTGAACATGC	TCGGGGGNGA	CGCGAATCCT	480
ATTTTATTTA	ATTGANGAAG	ATGATGTACA	AGCTGGTCAT	GCTGCATCAG	TAGGCCGTGT	540
TGGATCCCAG	ATCAACTTTT	ACTATTTAAT	GAGTCGTGGT	ATTTCTCAAA	GAGAAGCGGA	600
ACGCTCTGTT	ATACAGGGTT	TCTTAGATCC	AGTAGTACGT	GAATTACCTA	TCGAAGACGT	660
TAAACGTCAA	TTGAGAGAAG	TAATTGAACG	CAAAGTTTCT	AAATAATATT	TTGAAAATAA	720
AAGTTTGTA	TAGATATAGA	CTGTCGATAT	TGGTATAAGA	CTAATACAAC	GTCAGTATTT	780
AAATGATTAG	GATTTTATT	TAAGAAAGGT	CGTGAATGAA	GTGGCCGAAC	ACTCATTGTA	840
CGTTNATGAA	GTAATCAAGG	ATTTCCGAT	ATTAGATCAA	AAAGTCCATG	GCAAACGTTT	900
AGCATATCTT	GATTCAACAG	CGACAAGTCC	AACGCCTATG	CCAGTGTTAA	ATGTT	955

(2) INFORMATION FOR SEQ ID NO:150:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 462 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:150:

NAAATTTCTT	ATACGTTATA	CTTTAATTGT	TAATAAGCCA	CATAAATAAG	AGGGGGAAAT	60
GCTGTGTACA	AGCAAGGTGA	ACCAAATTTA	TGGACTGGAA	GGTTAGATAG	TGAAACAGAC	120
CCGAAAAAAT	TTAGACATTT	TCAAACAGTA	ACATTTGAAG	ATTTGTCTAA	GCTGGAAAAG	180
AGTAGTATGC	CATCAGGGGT	CGGTATATTA	GGCTATGCTG	TTGGACAAAG	GTGTTGCTTT	240
ANACAAGGGG	CGCATTGGTG	CAAAAGAAGG	ACCAGATGCG	ATTAAACAAG	CATTTGCAGG	300
TTGGCGGAT	TTGGATCACT	GTGAACTNT	AGTCGATTAC	GGAAATGTTT	NTCATGATCA	360

# EP 0 841 394 A2

TGAGGAATTA NCTNGATACN CAACCANGAA TTTGGTACTG TTTTGGCAAA TTGAGCTCTN. 420  
TCTTTAGAGC TTGGNNCATT GATAGGTTCT CTCGCAGGTG GT 462

5

(2) INFORMATION FOR SEQ ID NO:151:

(i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 752 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:151:

20

GGCACGAGCG	TAAGATTAAC	CATTGACCAT	AATATATATT	GTGTTTTTCC	AAAATCGGCT	60
CTGCTAATTT	TAAATAGGGG	CGATATATTG	TTATAAAACT	ATTGAAAAAT	TCTTGTGATA	120
GCATAGTGAC	ATCTCCTAAG	ACAAAATAGT	TAGCTTAGCT	AACCTTTTTA	CAACAATAGT	180
AATTATAAAA	CGGGAGCAAT	TAGAAATCAA	TATATAATTA	TTAAGAGCAA	AAATAATTAT	240
ACTTTGTAA	AATAAGCGTA	ATTACATGTA	AATAGGGGGA	TACTAATGAT	ATTGAAATGT	300
GCATCACATC	ACTCATTATA	TAGATCAGTT	AGATCGGTTT	AGTTCTCCAG	GAGATGTTAT	360
AAAATNACAT	TCAGGTGGGT	ATCATCATAA	ATATGGAACA	TTCAATAAAT	TAGGTTATAT	420
CAATGAAAT	TATATTGAGC	TACTGGATGT	AGAAAATAAT	GAAAAGTTGA	AAAAGATGGC	480
AAANACGATA	GAAGGTGGAG	TCGCTTGCGC	TACTCAAATT	GCACAAGAGA	AGTATGAGCA	540
AGGCTTTAAA	AATATGTGTG	TGCGNACAAA	TGATATAGAG	GCAGTTAAAA	ATAATCTACA	600
ANGTGAGCAG	GTTGANGTAG	TAGCGCCGAC	TCAAATGGAA	AGAGATACAC	ATAAAGATGG	660
TAAGGTAAAG	TGGCAATTGC	TTATATTATG	AATCAGGATG	ATGATGAAAT	TACGCACCAT	720
TTTTTATTCA	ATGGGAGAAA	GTGCCTCCAT	GC			752

40

(2) INFORMATION FOR SEQ ID NO:152:

(i) SEQUENCE CHARACTERISTICS:

45

- (A) LENGTH: 791 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:152:

CAATGTTTA	TCATGATATG	ATGAATATAA	TAATCGGGTA	TATAACTGTA	TGATTAATTA	60
CACAATAAAG	AGAGGGAATA	TAATATGNNG	NAAAGTGCAA	TTAAAGATGT	TGCTAGAGAA	120
GCTGGTGAT	CAGTTNCANC	TGTGTCACAT	ATTTTAAATC	ATAATGATAG	TCGTTTTTCC	180

# EP 0 841 394 A2

	GCANCACCGA TAAAAAACGT ACATGCTGTT CCAGAACGTT TAGGCTATGC CCCTAATAAA	240
	CATGCAAAAC AAGCTCGTGC GCGGCAGTAA AATTCAAAC ATTGGCGTCA TTTTACCTAG	300
5	CTTAACAAAT CCGTTTTTCT CAGCACTGAT GCAAAGTATT CATGACCATA AACCATCTGA	360
	TGTTGATTTA TGCTTTTTAA CATCTACAGC AACTGATNTG TATGACAATA TTAAACATTT	420
	AATTGATCGA GGTATTGACG GATTAATTAT CGCACAATAC ATATCATCCC CGGACGCCCT	480
	AAATAACTAT CTAAAGAAAC ATCATGTACC TTATGTCGTA CTGGATCAAA AATGACCATC	540
10	AAAGGCTATA CAGATTGTGN TCCGGACAAA ATTGAATATC AAGGGTGGGA CAACTTTGGC	600
	AGCACAAACA TTTAATAGAA CTCGGGTCAC AAACCATATT GATAATTGGT TGCAACCATA	660
	TTGACAATGA TGGNGAATAT TGTCGACTCC TGTCGCTGGA TTTGTCGATA CTTTGC GCGC	720
	GAATCAATTG CCAGAACCAC AAATCGTCCC TACTGAATTA TCTAAACGCC GTGGCTAACC	780
15	ATGTTGAATG A	791

## (2) INFORMATION FOR SEQ ID NO:153:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1314 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:153:

	TTTGTCTTTG GAATGGGCTC GTTTTTTTTA ACCTAATAAG AAATGATAGG GCATTTGAGA	60
	TTGGAAGGNC ATTTTGGCTT TGTGCAAATA ATACAATAAG CTAAATGTCT NTTTTTGT	120
35	TTGTGAAAT ATGATGGATG GCTTGTGTGG GCAAGTTTGC TAATTTAATA AGATATGCAT	180
	TTTTCAATTT AGGAGTTGGC CATGCATCTA CACTTTATAA TGGTGAGAGC GTGGTGAGGT	240
	ATTGTTAATC ACGCAATTGT AGCGAGGAGT TATTGCTACA TATGTCGTTA TGGCCTATTG	300
	ATTTTCTAAA ATAGCTGTAT CAGATCATGT GACNAAATAA AAATAATTTG TTGAAAGCCT	360
40	TTACATAACT TGTCTAGACA AGTTATACTC GTTTTAAGAC ATTAAGGGAG TGAAATATAT	420
	GGCTGTAAAA AGAGAAGATG TAAAAGCCAT CGTAAGNCGC TATTGGGGGA NAAGAAAATC	480
	NTGAAGCTGC ANCGCATTGT GTAACNCGAT TACGTTTNGT GCTTAANGAT GAAANCANAG	540
	TTGATAAAGA CGCATTAAGG AACAAACGCGT TGGTCAAGGG GCAGTTCAAA GCAGACCATC	600
45	AATATCAAAT TGTCAATTGGT CCAGGANCAG TCGATGAAGT GTATAAGCAG TTTATTGATG	660
	AAACAGGTGC TCAAGAAGCT TCGAAAGATG AAGCGAAACA AGCAGCTGCG AAAAAAGGGA	720
	ATCCAGTACA ACGTTTGATC AAATTGTTAG GGGAGATTTT TATACCAATA TTACCTGCGA	780
	TTGTGACAAC TGGTTTGTTA ATGGGGATTC AATAATTAC TTACAATGAA AGGTTTATTT	840
50	GGTCCCAAAA GCAC TTATTG AGATGTATCC CGCAAATTGC TGATATTTCA AACATCCATT	900
	AATGTGATTG CGAGTTACCG CATTTATTTT CTTACCANCA TTAATTGGTT GGAGTAATAT	960
	GCTGTGTATT TGGTGGTAGT CCGATCCTAG GCATAGTCTT AGGTTTGATT TTAATGCATC	1020
	CGCAATTAGT ATCTCAGTAT GATTTGGCAA AAGGGAATAT TCCGACGTGG AACTTATTTG	1080
55	GCTTAGAGAT TAAGCAGTTG AATTACCAAG GTCAAGTGTG GCCTGTTTTA ATTGCAGCTT	1140
	ATGTTCTAGC TAAATTTGA AAAAGGATTA AATAAAGTCG TTCACGATTC GATAAAAAATG	1200

# EP 0 841 394 A2

TTGGTCGTTG GACCGTAACG CTTTGTAGTTA CTGGATTTT AGCATTATT ATCATTGGAC 1260  
CAGTTGCATT ATTGATTGGT ACAGGTATTA CTCTGGTGT TACATTATA TTCC 1314

## (2) INFORMATION FOR SEQ ID NO:154:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 972 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:154:

ATAATTATTG NTGGAAAATA ACATTGAGCC CAATTCCCAA CAGTGGCCGA CATTTTCCAT 60  
CCATTTACCC CCCCCTGACT TTGTATCCGA ATTCNACCN TCCACCNAAG TGNCAATCAA 120  
TATGTTAATA ACCCAAATCC CAAATTGGAA ATATTTTGCC AAGCCAAGTT TTGATGATGC 180  
GCATTTGTCC TAATATTTTC CACTTAACAA GATCCGACCC CCAGCCAAGA AAATACATCA 240  
ACTGAGAAAT ATAATGNCCA AGACTGGTTT CCAACACACA TTCGACATAA TGAGTTGTCT 300  
AAATTGACTG AGCAACAACT TGTGATTGAG TTGCTTATGC ATATGATTCA TTATGGCACG 360  
AGCGTACATA TCATTGGAAC CCAAAGTATC TTAAATGATG ATAAAGTGAA TCAAGTATGC 420  
GACTATATCG AGTTACATTT TCATGAAGAT TTAAGCCTTT CAGAATTAAG CGAATACGTT 480  
GGGTGGTCAG AGAGCCATCT GTCTAAAAAG TTTACAGAAT CGCTAGGTGT AGGATTCCAA 540  
CATTTCTTAA ATACGACGCG AATTGAGCAT GCGAAACTCG ATTTAACATA CACAGATGAA 600  
ACGATTACTG ATATTGCATT GCAAAATGGC TTTTCAAGTG CAGCGAGCTT TGCGAGAACA 660  
TTTAAACACT TTACGCATCA AACGCCTAAA CAATATCGAG GTGATCGTCC AGCAATCACT 720  
GAAAATCAAC AATCGGCACA ACATAATTAT CACGACCGTG AATTGATATT ACTTTTAAAT 780  
GACTACATTG AAGAAATGAA TCAATTCAAT TGAAGATATT GAAAAGNTGA ACTTATAAAG 840  
AGATTGCCTT NAAACCAACT AATCAACAAC TAAATCCAAT TATAATCCAT ATTATCAAG 900  
TTGGGCTATT TGAGGAATTT GCTCCAATAC ACAGTATCAA TCCACAGTTG CTTACATGTT 960  
CATCCCATGA TT 972

## (2) INFORMATION FOR SEQ ID NO:155:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 503 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:155:



EP 0 841 394 A2

	AAACCATTTT	CCAAAACCAA	GCTGGGCATC	CAAGTTTTC	TATTGTAATC	TGCAGGTAAC	60
	CAAGGGAATG	CTTCCGTGAA	TTTAGGTGGT	AGCGTAACAT	CTATTCAACC	ATTACGTATT	120
	AATTTAACAA	GTAATGAGAA	TTTTACAGAT	AAAGATTGGC	AAATTACAGG	TATTCGCGT	180
5	ACATTACACA	TTGAAAACCTC	GACAAATAGA	ACTAATAATG	CTAGAGAACG	TAACATTGAA	240
	CTTGTTGGTA	ATTTATTACC	AGGGGATTAC	TTTGGTACGA	TACGTTTGG	ACGTAAAGAA	300
	CAATTATTTG	AAATTCGTGT	TANNCCACAT	NCACCACAAT	TACAACGACA	GCTGAGCAAT	360
	TANGAGGTCA	GGAATTACAA	AAGTGCCTGT	TAATATTTTCG	GGAATACCGT	TGGATCCATC	420
10	GGCATTGGTT	TATTTANTTG	CACCAACTAA	TCAACTACGA	ATGGTGGTAG	TGAGGCAGAT	480
	CAAATACCAT	CTGGTTATAC	CAT				503

(2) INFORMATION FOR SEQ ID NO:156:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1118 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:

	GGCACGAGAC	TCAAANCACT	GAAGCATTA	CAAAATAATA	CTATATTACT	GTCTAATCAT	60
30	AGACATGTTG	TATTTAACTA	ACAGTTCATT	AAAGTAGAAT	TTATTTCACT	TTCAATGAAC	120
	TGTTTTTTAT	TTACGTTTGA	CTAATTTACA	ACCTTTTCAA	TAGTAGTTTT	CATGCCACGA	180
	GCTATCCTAA	CCCACAGATT	AGTGATTTCT	ATACAATTCC	TCTTTTGTCT	TTACATTTTC	240
	TTAAATATT	TGCGATGTTG	AGTATAAATT	TTTGTTTTCT	TCCTACCTTT	TTCGTTATGA	300
35	TTAAAGTTAT	AAATATTATT	ATGTACAACG	ATTCAATCGC	TCTATTTTTC	AACTTTCAAC	360
	ATATTATTAA	TTCGGAAAGG	ACCACTTTAA	AATTTAACNG	GCCACAACAA	ATCAAATCAA	420
	TTAATCACTT	TTTCCAAAAT	AATCATATAA	GGAGGTTCTT	TTCATTATGA	ATATCATTGA	480
	GCAAAAATTT	TATGACAGTA	AAGCTTTTTT	CAATACACAA	CAAATAAAG	TTATTAGTTT	540
40	TAGAAAAGAT	CAATTAAAGA	AGTTAAGCAA	AGCTATTAAA	TCATACGAGA	GCGATATTTT	600
	AGAAGCACTA	TATACAGATT	TAGGAAAAAA	TAAAGGCACG	AAGCTTATGC	TACTGAAATT	660
	GGCATAACTT	TGAAAAGTAT	CAAAAATGCC	CGNAAGGAAC	TTAAAACTG	GACTAAAACA	720
45	AAAAATGTAG	ACACACCTTT	ATATTTATTT	CCAACAAAAA	GCTATATCAA	AAAAGAACCT	780
	TATGGAACAG	TTTTGATCAT	TGCACCATTT	AACTATCCTT	TTCAACTAGT	ATTGGAACCT	840
	TTAATCGGTG	CTATTGCAGC	AGGTAATACA	GCAATTATTA	AACCATCTGA	GTTGACACCA	900
	AATGTTGCAC	GAGTGATTAA	ACGATTAATC	AATGAAACAT	TTGATGCAAA	TTACATTGAA	960
50	GTTATTGAGG	GAGGAATTGA	AGAAACGCAA	ACGTTAATTC	ACTTACCTTT	TGACTATGTC	1020
	TCTTACAGGA	GTGAAATTGT	AGGCAAAATCG	TTTATCAAGC	TGCAGCGAAA	TTTAGTCCTG	1080
	TGACATAGAA	TGGTGGGGAA	ATCTCCAGTC	ATCGNNGG			1118

(2) INFORMATION FOR SEQ ID NO:157:

EP 0 841 394 A2

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 675 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:157:

CCANTTGCCG	TTTCTCCTA	AACACCAGCG	GNACGAGCTC	ATGATGGCAT	ACATTGTAAA	60
TCCGATAATT	GACAGTCCAG	TTGCTAATCC	ATCTAAACCA	TCTGTAAAT	TTACCGCATT	120
AGAAAAACCT	ACTTGCCAAA	AAACAATGAA	AATAACATAT	GCAAATGATA	GTGGGATTGC	180
TACATTCGTA	AATGGAATAT	GTATGCTCGT	AGAAAAATTC	ACCAAATGAA	ACACATTACT	240
TAAAACAAAG	AAAATAATCG	CAATACCAAT	TTGCGCCAAA	AACTTCTGTT	TACTTGTTAA	300
ACCTTGGTTA	TTCTTTTAA	CAACAATAAT	ATAATCATCT	ATAAAACCAA	TTAACCCAAA	360
ACCAATCGGT	CACAAATAAT	AACAGGTATG	ATTGGATTAG	CTTGATCTTA	CAAATATAAT	420
AGCCACCAAA	GACGGTTATC	ACAAATACTT	TAATAGAAAT	GGTTAGGCCA	CCCATCGTTG	480
GTGTACCACT	CTTCTTCATA	TGGCTTTGTG	GACCTTCTTC	TCGAATACTT	TGACCAAATT	540
TCATCCTTTT	TAATGTAGGT	ATTAAAACAG	GTACCAAAAC	AAATGTAATC	ACTAGCGCTA	600
ATAACGCATA	TACAAAAATC	ATAACTATCT	CCTCTTCTTA	ATCCAGACTT	TTTTAACCAC	660
TAATATATTA	TCAAG					675

(2) INFORMATION FOR SEQ ID NO:158:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 746 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:158:

CTTAGATACA	ATTACTCAAT	ATGATGTACT	GGAAGCANTT	ATAGATACTA	AAAAACACAT	60
TGNTGTNCG	ATGANTNTCA	TCTTCACATG	ANNTANCGGG	TTGATTAAACN	AAGATTGCAG	120
ACCGTGNTGT	TGTGATGANA	AATGGTCANC	TGATAGAGCA	TGGTACACGT	GAATCAGTCT	180
TGCATCATCC	AGAACATGTT	TATACGAAGT	ATTTATTATC	ANCGNAGAAG	AAGANTAATG	240
ATCATTTTAA	ACATGTGATG	AGGGGTGATG	TACATGANTA	AAGTTACAGA	TGTTGAAAAA	300
TCATATCAAA	GCNCACATGT	TTTTAAGCGT	CGTCGAACAC	CTATCGTGAA	AGGTGTGTCA	360
TTTGAGTGTC	CAATCGGTGC	GACGATTGCG	ATTATCGGAG	AAAGTGGTAG	CGGTAAATCG	420
ACGTTGAGTC	GTATGATATT	AGGTATTGAG	AAACCGGATA	AAGGCTGTGT	AACCTTAAAT	480

EP 0 841 394 A2

GATCTACCGA TGCATAAGAA GAAAGTCAGA CGTCATCAAA TTGGTGCTGT ATTTCAAGAT 540  
TATACGTCAT CATTACACCC ATTTCAGACT GTTAGAGAAA TCTTATTTGA AGTGATGTGT 600  
CAATGTGATG GACAACCTAA AGAAGTTATG GAAGTCCAAG CAATTACATT GTTGGGAAGAA 660  
GTCGGTCTAT CTAAGGCATA CATGGATAAA TATCCTAATA TGTATCAGG TGGAGAAGCG 720  
CAGCGTGTG CGATTGCGCT CGTGCC 746

(2) INFORMATION FOR SEQ ID NO:159:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 692 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:159:

AATGTCAGAC AGATACTGCC ACAAGATGCG TGCNTATGAG ATTTGCTGT GTATGAATAG 60  
CGACACGAGC GGCACGAGCG CTCACNTCAT TTCCAATTAA AACTAATGCC TAAATCTGAT 120  
GCAGTAAAT CTATCATGAT TCACTCTTAA CATCCGTATT TCGTGCTACT AATTGATGTC 180  
TTGCATTGAA AAATTGACCA TAGCTTAAAT ATGTCGCAAT CAAAGCAGAC ATAATGNTCG 240  
CAGTTGTATG AATAAACACG ACTAACAATT GAAATTTAAT CGCTTGTAAG GGTGGTACGC 300  
CACCAATAAT TAAGCCTGTC ATCATAACCAG GAATCGACAC AAGCCCATAT GTTTTAACCG 360  
AATCAATTGT TGGCACCTAT AGCTAAACGA ATACTTTCAC GTATTGCACC TTTAGAAGCC 420  
AATTTAGGTG TAGCTGCAAG TGATAATTTA GATTCAATAT TAGTACCATC TTGTACGAAT 480  
GCACGATCTA AATTCTGGTA AGCTAAATTA ATTGCAATCA AGCCATTATT TNCAAGCATN 540  
CCGCCGATAG GTATAACTTC ATTGGCTGTA AAATGAATTG CCCCTGTAGC TACAGTACCT 600  
GCAAGTGGA ATGCTGTTCC AATGAAGATN GCTGGAAATG GTNTCCAAAA CACANGGGGC 660  
ATCACTGTGA TGCTCGACTA ATGGTAAGAG TC 692

(2) INFORMATION FOR SEQ ID NO:160:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 857 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:160:

AATATAGCGT TTNACCCCT TTGTGTNACC CTCCCCAAGA GATATAACAT NCCGCCGTNT 60

EP 0 841 394 A2

5 AAAATCAATT AGAAATATCT TTTTATTCTG ATAATAGACA CAGTATAGAC ACATTTTAAAT 120  
GGGTGATACC ACTTGTAATA TCACGGGGTT GTNATGTNTT GNATATCANT NAAATACTTA 180  
TATANAAATA TTGCTCGGAA TATAAAAAGN TAAATAGGNT TTTGAGTTTT AAATATGAAA 240  
TACAAAGCGC CCANTCGAAC AAAGTATTTA TATTAAAATA TGGAAAATCC ATCANTATTA 300  
AATTAAAATN GTTTTATTAT GATAAAGTGA AAGTAGGTAA GTCTATGGAA GGTCTTAATC 360  
ATCGAAGAAA TACAGAAAAA GAAGAGACAA CACAAACGCA ATCAGTTGCA CCTAATACAG 420  
10 GTGAAGAGGG GATGTCATCA GGCAAGTAAC ACAATCAANT AAGACGTCCG ACATACATAA 480  
ATGAATCTAT CAATAAACAA ATGGAAGCCA AAGCGCATGA AACAGCGCAA AATGCAGATT 540  
TAAAAACCGA AGCAAGAAGT TTATTTGATA ATGCAACCAA ATCAATCGGG AGACTAGCCG 600  
GCAATGATGA AAGCATAAAT CTTAATTTAA AAGATATGTT TTCTGAAGTA TTTAAGCCGC 660  
15 ATACTAAAAA CGAAGCAGAT GAAATATTTA TAGCGGGTAC TGCTAAAAC TACGCCAGCAA 720  
TTTGTGACAT ATCAGAAGAA TGGGGGAAGC CATGGCTCTT TTCTCGAGTA TTCATCGCTT 780  
TCACAGTAAC ATTTATTGGA TTATGGGTCA TGGCAGCGAT TTTTAATAAC AATGACGCTT 840  
GTACCGGTGC TCGTGCC 857

(2) INFORMATION FOR SEQ ID NO:161:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 907 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:161:

35 CTCACCTTGT TTTCCAGTAT GGAACGCTTG CCCTAAGTCC TTAATTGCAT TATAAAATTC 60  
AGGCGCTAAA ATAATCGCAA TTGCCGCGAGT TTTAAAATCA ATATTATGAA ATACTACTAA 120  
GCTTAGCGTT GCTTCCAATG CAACCAATCC AATACCTAAC ATACTTATAA ATTGAGCAT 180  
40 TAATCCCGAT AAAAAAGCAC TCGCTAAAAT GCGCATTGTT AAAGTTCTAA ACTGAGTACT 240  
ATCGTCGTAA ATATGCTTCT CTGTTTGCTC TGTACGATTA AATAGCTTTA ACGTCACTAA 300  
ACCTNTAGCA ATATTTAAAA ACCGNCGACT AAATTGATTC ANATAAGTCA TTTGATCTTT 360  
TTGACGCATC GAGCGTTTTC AAACCGAAAA TAATATAAAA CAAAGGAATA AATGGTGCAG 420  
45 NTATTAACAT AATTAATGCG GNATTGAAAT GGATGAAAAA CAATGCAATG ATTATGATGA 480  
GCNGAACCAT CAATCGATTT GAACAACCTG AGGCAAATAA CTCTTATAAA AAGGTGCTAA 540  
ACCATCAATG TTTTCTGTGA GTATAGTCAT TTGTTACCG ATTGGATGAC CATTATTTTT 600  
ATAAATAACC CGCTGTCTAA GCATATGCTT AACTTTAAAT GCTAATGTAT CACCTAACCA 660  
50 TTGATTTAGA AATTGCACAG NTGCTCTTAA AAGTAAAAA CCTAATAAAA TAAATAATAC 720  
AATCCATAAA CCTTGAAATT GATGTCTTAT AATTTTAGCT AAAAAATCTG CTATTAAAAAT 780  
ATTGTGCGTT ATAACGAGTA TGCCCGAGAC CAGTACTGAC CAAGAACATG AGTACCGGAA 840  
AAATTTTATA TTGAAACAGT ATTGTTGTTA ATTTTTCACA ATTATATCAC CTAACCTATA 900  
55 TAAAGTT 907

# EP 0 841 394 A2

## (2) INFORMATION FOR SEQ ID NO:162:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 774 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

AATTTTAAAG	ACCCCNCGCA	TAAATANCCA	TCCCACCTAC	TTATCCAAAA	GTTANGTGGA	60
TGGTTTTTCA	ATTAAAATTA	ATATTAGTGT	AANCCAATCA	AAGATTTAAT	CNAATATGCC	120
CCTGCTCAAA	ACATTTCTCT	ATTTAATTTG	CTTTACTTTC	AATTTAATAT	CATTATCCAC	180
AACACTTGGC	GTGTCACTCG	TATTATTTTC	CATCTTTGAC	ACGTTTATCA	TCATTAGGAN	240
TCGGCACCAG	ATAAAATTGC	GATAAATGCC	ATGATTCCCA	TTAATACGTT	AACCCAAAGT	300
GCAATCATCG	CACCTGTATG	AATGCTCGTT	GCAGCAACTG	CACCAACATA	TACAGCACCA	360
CTAATTGCGA	CACCGAATGC	GCCACCAAGT	GATGAAGCCA	TTTATAAAT	ACCTGAAGCA	420
ACGCCAACTT	TATCTAACGG	TGCATTGCGA	ATAGCTGTAT	CTGTAGAAGG	TGTTGCATAA	480
ATACCTAAGC	CTAGTCCGAA	ACATAAATAT	CCTACGACAC	AACTGATAAC	ATAAAATATG	540
CCTGGTAAGA	ATACTANTGA	AATAAGTGCA	ATNCCAATGA	CCACAATGNA	TGTACCTNAT	600
AACATTGGTC	GCTTAGAACC	CANTTNTGGT	NATAATAATT	TTTCACCAAC	TCGAATCATC	660
AATAACAACA	TGATTAAATA	AGTAANTGAT	NAGTATCCTG	CCTGCCATNC	TGTATAACCT	720
AAACCTTGTT	GCACGCATGT	ATTCGCTACA	ATTNATGTAC	CTACAACNCC	GTTG	774

## (2) INFORMATION FOR SEQ ID NO:163:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 773 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:163:

CTCTGATTCA	ACAAAATGAT	TATTCTTTAC	GTAAACTACC	TTTTTTTATT	TGAGATGAAG	60
CATATGCTTT	TAATAATATT	GTCCCAATAA	TACCAACTGA	AATAATATTT	AATACTGCAG	120
AGATAACACC	TTGTGTATAA	ACCTTGTTAG	CCGGTTCGTT	ATAAATCAAA	ATATCTAATG	180
TTGGTGCAAT	AAGTGCCCAG	CAAATAATAT	TCGCAATAAT	TTGACCGATA	TTAAAAATAA	240
TCATCGATTT	CCTAGAAAAT	AGGCATGAAG	AAAGATTTAA	TTTTAGGGCC	AATCCATCCA	300
TATTAAACAG	GCGATAATTC	CTGAACAAAT	AACCCAACCT	CACCAAGCAC	TACCCGTATG	360

# EP 0 841 394 A2

TCGGGGAAAT CTTTAATAGC GTGNCCAACT AATCCAGGCA TTAAACCAGC AAAAGGCCCA 420  
 AATATTGCAG ATATTAATGC TAAAAATGCA TAAGATGTTT CTATATTCGT ATTAGGAAAA 480  
 5 CCTGTTGGTA TTACAACAAA ACGCCCTAAA ATCACAATA CCGCNGCTCC TATACCAATC 540  
 GCAACAACAG TTTTAACTGA AATATCNTGT TTTTTCATCT TCATTACTCC TTACATAAAA 600  
 AATTCATTAA ATTGATGGTG CTTTAGATAA ATGAATCGTC CAATCATTTC CAGTACCAAT 660  
 ATGATATAAA TCTGAAAATG AGTCTCGATT GACTGCTACA CCAATATTTA CTAGCGAGTT 720  
 10 AACATACACA AGAGGTTTAC CCACATTAAC ATCTGCAAAC GATCGCTCGT GCC 773

## (2) INFORMATION FOR SEQ ID NO:164:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 676 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- 20 (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:164:

CATAAGACAT GTAAATTCTA AGATGATTG TTGGATAGGG TAGCTCCATA TAATTCCTGA 60  
 ATCCNATCAT TATTACATTA ATAAAAAAA CACCACAAT TGTGGGTGAT TGTAATGAAC 120  
 30 GTATTATCTT GGCACGAGTA CTCAACGATT AATTGTTTCGT TAATTCAGC AGGTAATTCG 180  
 CTACGTTCTG GTAAACGTAC GAAAGTACCA GTTAAGCTGT CAGCATCAA GTTTAAGTAC 240  
 TCAGGTACGA AATTGTTGAT TTCAACTGAT TCAACGATGA TGTTTAGTTT TTGAGATTTT 300  
 TCACGAACTG AAATTGTTTG ACCAGGTTTA ACAGAATAAG ATGGAATATC AACACGTTTA 360  
 35 CCATCTACTA AGATATGACC GTGGTTAACT AATTGACGTG CTTGACGACG AGTACGAGCT 420  
 AAACCTAATG AATAACAAC AGCGTCTAAA CGACTTGCTA ATAAAATCAT GAAGTTTTC 480  
 CCCGCGTACA CCCNAATTTT TTACCAGCGA TGTCAAATGT GTTACGGAAT TGTCTTTCAG 540  
 TCAATCCATA TTANGTAACG TAATTTTTGT TTTTCACGTA ATTGTAAACC ATATTCTGAT 600  
 40 AATTTTTTAC GTTGGTTTGG ACCATGTTGT CCTGGTGCCT AAGGACGTTT TTCTAATTCT 660  
 TTACCAGTCT CGTGCC 676

## (2) INFORMATION FOR SEQ ID NO:165:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 397 base pairs
- (B) TYPE: nucleic acid
- 50 (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:165:

# EP 0 841 394 A2

GGCACGAGCG ATGATGGCTG TCGGAACAGG TGCATTGGT GCGCATGGTT TACAAGGGAA 60  
 AAATAAGTGA TCACTATTTA TCAGTATGGG AAAAAGCAAC GACGTATCAA ATGTACCATG 120  
 GCTTAGCATT ATTAATTATA GGTGTAATTA GTGGTACAAC TTCAATCAAT GTTAACTGGG 180  
 5 CTGGCTGGTT AATATTGCT GGTATTATTT TCTTTAGTGG ATCATTATAT ATTTTAGTAT 240  
 TAACTCAAAT TAAAGTTTGA GGTGCGATTA CGCCAATTGG TGGCGTATTG TTCATCATTG 300  
 GCTGGATAAT GTTAATCATT GCGACATTCA AATTGCTGG TTAAATTTTA AAACCTTAGA 360  
 TTACCTATGT AACTAAACAT TAAATTTTTA ANAAAAA 397  
 10

## (2) INFORMATION FOR SEQ ID NO:166:

### (i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 739 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 20 (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:166:

25 TATCTTTAAG AACAAATTGAA GCCGCTAACC ATCGGAGAAA GCTGTAAAAT CCAAGTGTTG 60  
 GCCTGCTAAA TACCCACAAT CATCACGGGT CGTTGCCTTG TACCACAATA GAAGGCAGCA 120  
 TTATCACCCA AATATTGCA TAGCTAATTG TGATAAACT GTCGTTCCG TTTGTGGCAT 180  
 30 AATTCCATAA ACATATGCTA AACCACCGAT ACCAACTAAT AAAACGCTA AAATTGAACC 240  
 CATAGCAATT AACGTTTTTA CAGCATTATT AGCACTTGGT TCTCTAAAAA TTGGTGACCG 300  
 CATTTGAAAT AGCTTCAACA CCTGTTAATG ATGAAGCCCC TGATGAAAAA GCTCTTAATA 360  
 GCAAGAATAA TGTTACTCCA GGANCCGCG TCCCTACTGA TGCATGCATA TGTGGTTGAA 420  
 35 TATCTCCTGT CGCCACACGG AAAGTACCCT ATAAATATTA ATATCACTAA CCCCATATG 480  
 ANAAGATATA CTGGATAGGA TAATACGGTG NCAGATTCAG TTAAACCCAC GTAAATTTAA 540  
 TATTAAAATA AAAAGTACAA GTAAACATGC AATCAGTACT TTATGCCCAT ATAACTTGG 600  
 GAATGCAGCA ACANATGCAT CAGCACCAGA TGATATACTA ACAGCGACAG TCAGTATGTA 660  
 40 ATCGACTAAT AATGAGCCTC CTGCAAGCAA TCCCCATTTT TCTCCTAAAT TGGTCTTGGA 720  
 CACCATATAC GCTCGTGCC 739

## (2) INFORMATION FOR SEQ ID NO:167:

### (i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 507 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:167:

5	GGCACGAGCG AGAACGATTG AAGCTACAAT ACCTGATGTT GCTGCGGGAA GTACGACTTT	60
	AGTTGCTACT TCTAATTTAG TTGCTCCAAG TCCATAGGCA CCTTCTCGAA TTTTATTTGG	120
	TACAGATGCC ATTGCATCCT CACTCAAACCT TGTGATGAGA GGGACAATCA TAATACCGAC	180
	AACTAAGCCG GGACTTATAG CATTAAACTC TCCAAGACCT GATATGAAAG ATCTTAATAC	240
	TGGTGTAACA AAGGTTAATG CAAAGAAACC AAACACAATT GTTGGTATTC CTGCTAAAAT	300
10	TTCTAATATC GGTTTAATTA TGCCTCGGGC ACGGGCACTT GCATATTCAC TTAAATAAAT	360
	TGCTGCACCA AGCCCGACTG GAACTGCAAA TATAGTCGCA ATAACGTGA TTTTTAAAGT	420
	CCCTATTATC AATGCCAGAT ACCAACTTA GGGTCTGAAC CGGTAGGATT CCAGTAGTAG	480
15	AAATAGAAAT CAGTATTGGA ATTCTGG	507

(2) INFORMATION FOR SEQ ID NO:168:

(i) SEQUENCE CHARACTERISTICS:

20	(A) LENGTH: 753 base pairs
	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: single
25	(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:168:

30	GAATAGTCTA CTCATTCATC GTCCTTCACC ATTGATGGAT CCAGAACAAG TTGCTGATGC	60
	ATTAATAAAA CTTGTTTAAAC AAGGTAAGTT GAAGTCATTC GGGGTGTCGA ATTTTAATCA	120
	TTCACAATAC CAATTGTAA ATCAATATAT TATGAAAGAA AGACTACATA TTAGCATCAA	180
35	TCAATTAGAA TTATCGCCAT ATCACGTTGA TAGTTTACAA GATGGAACAA TGGATTCAAT	240
	GTATCAAAAC CATGTTCAAA TTATGGCTTG GAGTCCTTTT GCAGGCGGTA AAATTTTCGA	300
	CAAGGAAGAT ATTAAAGCGC AACGTATTAT GAAAGTTGTC AATCAATAGC TGACAAATAT	360
40	GGTGTGAGTG ACACAGCTGT GATGATAGCA TGGTTAGTAA AAATACCGCA TAGTACCATG	420
	CCGATACTGG GAACAAGTCA GTTAAAGCGT ATNGATCAAG CAATCGAAGG GCTACAACCTT	480
	AATTTAGATG ATCAAGTCGT GGGTTGACAT TTACAACGCT ATTATCGGAC AAGATATCCC	540
	GTAAACTTAN NNACNCAAT ATCATAAATG GAGCATACCA TGACAAACGA AGATAAACGT	600
45	TTCGAACAAT TAAGATTGTA ACGCAAATTT ATAGTTATCC CGTATTTAAT TTATGCAGTC	660
	ATTGTATTAC TATTAAATAT TTTCTATTCT GATTGAAAA TAACAATGAC ATTATTCGGA	720
	CTTTTCTTTG CGTATAATGT AGTCATTTTG TTC	753

(2) INFORMATION FOR SEQ ID NO:169:

(i) SEQUENCE CHARACTERISTICS:

55	(A) LENGTH: 542 base pairs
	(B) TYPE: nucleic acid



# EP 0 841 394 A2

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:169:

10	CCCATTATTA	TTTNGAAAAT	GAATTGGAAA	ATACAGTAGA	CGAAGTGTGG	GTTGTATACA	60
	CTTCTGAAAG	TATACAAATG	GATCGTTTAA	TGCAACGTAA	TAATCTGNCA	TTAGAAGATG	120
	CGAAAGCACG	TGTCTACATA	CATCAACGCG	TACGTAAAGT	AATGACTTAT	TTATAGTGTA	180
	ATATTAATCT	TCTTCTCCGN	AATTCGGNTT	TNCAATATAA	CCTTCTTCTT	CTAACAANCT	240
15	CTCAAGGTTG	TGNGTNAATN	CAAGTTTATC	CCCTAAATTA	TCGATAACAT	GATCGGCCAN	300
	TCGGNTNNNN	NNATCAATAG	AAATTGGNT	TATAGACACG	TGCTTTCGCA	TCTTCTAATG	360
	ACAGATCATN	ACGTNGCATT	AAACGATCCA	TTGTATACT	TTCAGAAGTG	TATACAACCC	420
	ACAGTTGCTC	TACTGTATTT	TCCAATTCAT	TTTCAAATAA	TAATGGAATA	TCCACGAACC	480
20	ACATTATATC	CTTGGTTTTA	AATATTCTTG	CTTTTCTTCT	GCCATAATAT	CTCGCACGAA	540
	TA						542

(2) INFORMATION FOR SEQ ID NO:170:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 731 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:170:

40	TCAATCCANC	CTCCAATTGC	TGATGAAGAA	CCTCCTTCAG	CNCCCACCCA	TTANNGGGGC	60
	CANAGTCATA	AGTAACAAC	TTTGAACCAA	TTGTATCTTC	AATTGAATCT	GTAATCTTAT	120
	CTCCCGCTTC	TTCCCATCCT	AAATGTTCTA	ACATTAATAC	AGAACTTAAA	ATTACTGAAG	180
	NTGGATTAC	TTTATTTAAA	CCTGCATATT	TTGGANCTGA	GACCATGTGT	TGCTTCAAAA	240
45	ATAGCATGAC	CTGTTTCATA	ATTAATGTTT	GCACCTGGCG	CAATACCAAT	NCCACCAACT	300
	TGTGCAGCTA	AAGCATCTGA	AATATAGTCA	CCATTCAAGT	TCATAGTTGC	TACAACATCA	360
	TGCTCAGCTG	GACGAGTTAA	AATTTGTTGT	AAGAAAATGT	CAGCAATAGA	ATCTTTAATG	420
	ATAATCTTGC	CTTCTTTCAC	AGCTTTTCT	TGAGCAGCAT	TAGCAGCATC	TCTGCCTTCA	480
50	TTTACAACAA	TTTCGACATA	TTGTTGCCAA	GTGAATACTT	GCATCACCAA	ATTCAGATAA	540
	TGCTAAATCG	TAACCCCACT	GCTTAAATGA	GCCTTCTGTA	AATTTCATAA	TATTACCTTT	600
	ATGAACCTAA	AGTAACTTGA	TTTACGGGTT	ATTTATCGAT	AGCATATTGG	TATAGCTGCT	660
	CTACTAATCG	CTCAAGTCCT	TCTTTTAGAA	CTTGGTTTAA	TACCAATACC	TGAAGGTTCT	720
55	TGGAATCGAT	T					731

# EP 0 841 394 A2

## (2) INFORMATION FOR SEQ ID NO:171:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 695 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:171:

TTTGACATCT	CTCAAAAATT	AAAGCATAAA	GGTTTATCGG	AAAGGCGCAA	AATTCACAGT	60
TGTTTGTGG	TTTGTTATTT	CCCTCNCAAT	ATTCAGTATT	AGACATTTAT	AGTTTGGA	120
ATGCGTGATA	ATTAGTTGTA	TTCAGTTATT	AAGTAATAAA	TTTTTGGAGG	CAGAACATCA	180
TGAAATTAAC	ATTAATGAAA	TTTTTTGTGG	GGGGATTGTC	AGTATTATTA	AGTTATATTG	240
TATCTGATAA	CAATAACCTT	GGGAAAGAAT	TTGGCGGTAT	ATTGCAACG	TTCCGGGCA	300
GTATTTTATG	TGTCTATGTT	TATTACAGGT	ATGCAATATG	GTGATAAAGT	CGCTGTGCAT	360
GTAAGTCGTG	GCGCAGTGTT	TGGTATGACA	GGGGTATTAG	TTGTATTTTA	GTAACATGGA	420
TGATGTTACA	TATGACGCAC	ATGTGGTTGA	TTAGCATTAT	TGTTGGTTTC	CTAAGCTGGT	480
TCATCAGTGC	AGTATGTATT	TTTGAAGCGG	TAGAATTTAT	AGCACAAAAA	AGATTAGAAA	540
AGCATAGTTG	GAAAGCTGGA	AAATCGAATA	GTAAATAGTG	TGAACGTAAT	CTCTTAACCTA	600
GGACTAACTT	TGCAAGCATT	GAATAGCATG	GAAAAGTTGC	ATCATTAAATA	AGTGAAATTC	660
AAGTTGGCAT	TGAGAAAATT	ACAAGCGCTC	GTGCC			695

## (2) INFORMATION FOR SEQ ID NO:172:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 612 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:172:

GGCACGAGCT	CGGCTTGACG	GTAGGAAATA	TCAGCACGAG	CTTGATTCAA	CANCCGAATC	60
AGGGAGGAAA	TGCAGGTATA	ACTCAATCTA	AGTTCGCTAA	GAGACATCCA	ACGTTTTACN	120
AGGCAAATCC	AAGCNAAAAG	GATTGCTTCN	AATAATATCC	CACCAAACAA	TTTAAGACAT	180
TATGCTGTTA	AGAGGTCACC	AACNATATAT	ATAGTGGTTA	CGGATCAGTT	ATTAGCGTTC	240
TTTAACAACA	GATATTGGCG	CTCACAGTTT	AACCCAAGAG	GTGGTTGGTC	TCCAAGTGGT	300
CCAAGAAGAT	ATGCGAATGG	TGGTTTGATT	ACAAAGCATC	AACTTNTCTGA	AGTGGGTGAA	360
GGAGATAAAC	AGGAGATGGT	TATCCCTTTA	ACTAGACGTA	AACGAGCAAT	TCAATTAACCT	420

# EP 0 841 394 A2

GAACAGGTTA TGCGCATCAT CGGTATGGAT GGCAANCCAA ATAACATCAC TGTAATAAT 480  
GATACTTCTA CAGTTGAAAA ATTGTTTGAA ACAAATTGTT ATGTTAAGTG ATAAAGGAAA 540  
TAAATTAACC GATGCGTTGA TCAAACGTCT CTTCTCAGGA TAATACTTAG TTCTATGATG 600  
CACTTAGAAG TT 612

## (2) INFORMATION FOR SEQ ID NO:173:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 605 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:173:

GGCACGAGCG ACTTTTCTA GGCATAATCG AATTGACAAT GGTACTCAAG CTAAAAATG 60  
GCCCACTTAA TTCAGGCAAT AACAGACTAG GCATAACATT ATTTTTCATC AATTTAAATG 120  
TGTA AACAT CGATGACATT GTCTGTTGCT GTTGTCGATA AACATTCATA TCGTAGCGGT 180  
CTGCAAAATC TTTAATGCGA TATGCCGCGC GCACGAGACA TGACAGGTAA TGAATCATGT 240  
TTGAATTGTT CGTCTACGGC ATCTTTTGA ATAGGTAATC CAAAGAAACC TGCAATACCA 300  
ATCGTTTCAA AGGGCCCTGC TGCTTCGATA TGTCTACGAA ATGGTTCTGA ACGAACATCT 360  
ATACAAAATG CAATTGCGC TTTCGTTGAT GTGCCACCT GATTTAGCTC GCTATTATTT 420  
TCATCAACTG CTTGTGTGTC AATTAACAAT ACTGAATGTG GCTGATTAGC GTTATCATTT 480  
TCTGAGACAT TTACTTGGTT TACATCTAAT GCGCCCGCCA CACTTTCATG ACTGCTTTAA 540  
TTTTTGTTTT AACTGAGATC GTATGTCAAT TTCCAGGCA ATTAGCCATA AATTNTTAA 600  
TACAT 605

## (2) INFORMATION FOR SEQ ID NO:174:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 508 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:174:

GGCACGAGCT GATCAAGGGT GTAAGTTGGT AGTGGTCAAT AAAGAACAAT CATTACCAGC 60  
TAACGTAA CAAGTGGTTG TGCCGGACAC ATTAAGGAGT AGCTAGTATT TCTAGCAACA 120  
ACAACATTAT ATGGATTATC CGGAGTCATC AGTTAGTGAC ATTTGGTGTA ACGGGTACAA 180

EP 0 841 394 A2

ATGGTAAAAC TTCTATTGCG ACGGATGATT CATTTAATTC AANGAAAGTT ACAAAAAAAT 240  
 AGTGCATATT TAGGAACATA TGGTTTCCAA ATTAATGAAA CAAAGACAAA AGGTGCAAAT 300  
 5 ACGACACCAG AAAACAGTTT CTTTAACTAA GAAAATTAAA GAAGCAGTTG ATGCAGGCGC 360  
 TGAATCTATG ACATTAGAAG TATCAAGCCA TGGCTTAGTA TTAGGACGAC TCGGAGGCGT 420  
 TGAATTTGAC GTTGCAATAT TTTCAAATTT AACACAAGAC CATTTAGATT TTCATGGCAC 480  
 AATGGAAGCA TACGGACACG CTCGTGCC 508

(2) INFORMATION FOR SEQ ID NO:175:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 750 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:175:

GGCNCGAGAC TGTCGAATAT TTAGTAGTAA CTTTCAAGTA CAAGCGTATG ACTTATCGAC 60  
 CGAACGGTAC AAATAAAGTA TTTGTTAAAA GAAAAGAAGC GGGNTCATGG TCTGAGTGGT 120  
 CAGAAATAGC TATTAATGAT TACAATACAC CTTGTGAAAC TGNTCAAAGT GCCCANTCAA 180  
 30 AAGCTAATAT GGCCGAAAGT AACGCTAAAT TATACGAGA TGACAAGTTT AATAAAAGGC 240  
 ATTCGAGTTA TTTGTGGATG GAACAGCAAA TGGTGTGGGC TCTACATTGN ACTTAAATGA 300  
 GAGTTTAGAC CAATTTATTT TATTAATTTT TTATGGGACT TTTCCAGGTG GTGACTTTAC 360  
 AGAGTTTGGC AGTCCTTTTG GAGGAGGAAA GATTTTCATTG AATCCCTCAA ATCTTCCAGA 420  
 35 TGGTGTGGA AATGGTGGAG GTGTTTATGA GTTTGGATTA ACTAAATCTA GTCGTACATC 480  
 TTTAACTATA TCAAACGATG TCTATTTCTGA CTTAGGAAGT CAAAGAGGCT CTGGTGCGAA 540  
 CGCAAAATAGA GGGACAATTA ACAAATTAT AGGAGTGAGA AAATAATGCA AATATTAGTT 600  
 AACAAAGCGTA ATGAGATAAT TTCATACGCT ATCATTGGTG GCTTTGAAGA AGTATGATAT 660  
 40 TGAAATTACA GAAATTCTCT CAAGTTTCTAG ACTAAGCTTT AAATATCAAT GGAATAGTT 720  
 TTACGAAGAT ATCCGAGAAA AGATGACTGC 750

(2) INFORMATION FOR SEQ ID NO:176:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 787 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:176:

EP 0 841 394 A2

	AAATGCCGGG	GGAGCTCAAG	TATATGACTG	AAATAACATT	CAAAGGTGGA	CCAATCCACT	60
	AAAAAGGTCA	ACAAATTAAT	GAAGGTGATT	TTGCACCTGA	TTTTACAGTG	TTAGATAATG	120
5	ACTTAAATCA	AGTAACATTA	GCAGATTATG	CTGGTAAAAA	GAAATTAATT	AGTGTGGTAC	180
	CATCAATTGA	TACAGGTGTT	TGTGATCAGC	AGACTCGCAA	ATTCAACTCT	GAAGCTTCTA	240
	AAGAAGAGGG	GATTGTGCTT	ACAATTTTCAG	CAGACTTACC	ATTCGCACAA	AAAAGATGGT	300
	GCGCTTCAGC	AGGTTTAGAC	AATGTCATTA	CATTAAGTGA	CCACCGTGAC	TTATCATTTG	360
10	GTGAAACTA	TGGCGTTGTT	ATGGAACGAA	CTTCGTGCCG	AATTCGGCAC	GAGCTCGTGC	420
	AGTATTTGTA	TTAGATGTAG	ATAATAAAGT	TGTTTATAAA	GAAATCGTTA	GTGAAGGTAC	480
	TCGATGCCCC	AGATTTTGAT	GCTGCTTTAG	CTGCATACAA	AAATATTTAA	TCATTAAAGA	540
	GATAAATCTT	AAAATGTATA	CATCGTGTCC	ATCGTTGTCA	ACAGCATTAA	AATAGAATTG	600
15	TTTTCTATGA	TTGCTAAGAC	CTATGGGCAC	TTTTTATTGG	AGAGGGACGA	ATATGGCAGA	660
	ACAACAAACA	ATTATGGAAC	GCTTGTTTCA	TACATTAGAT	GAAAAAGCTA	AAACATTAAA	720
	TAATGAAAAT	GGCCAAAGTT	TTATTGAAAA	TCTTGGGCTA	GCAATGGAAC	AAGTATATAC	780
	CAATGAA						787

(2) INFORMATION FOR SEQ ID NO:177:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 568 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:177:

35	GGCACGAGCG	CGTCATATAA	TGAAAGTAAT	GATAAAAAGA	AAGGATAACT	TAATGTGAGT	60
	CAAGAACGTT	ATTCAAGGCA	AATTTTATTT	AAACAAATAG	GTGAAATAGG	TCAAAGCAAA	120
	ATAAATCAAA	AATGTGCGTT	GATTATTGGT	ATGGGCGCAT	TAGGTACACA	TGTGGCCGAA	180
40	GGACTTGTTA	GAGCAGGCAT	TGCTAAACTA	ATCATTGTTG	ATAGAGATTA	TATTGAATTT	240
	AGTAATTTAC	AAAGACAAAC	ATTGTTTACT	GAAGAAGATG	CTTTGAAAAT	GATGCCTAAA	300
	GTGGTTGCAG	CTAAAAAGCA	TTTGCTAGCG	TTACGTAGTG	ATGTTGATAT	TGATGGTTGT	360
	ATTGCCCATG	TGGATTATTA	TTTTTTGCGA	AACACATGGA	CAGGACGTTG	GACGTTATTA	420
45	TTGATGCAAC	CGATAACTTT	GAAACACGAC	AACTGATTAA	TGATTTTGCA	TATAAACATC	480
	GTATTACCTG	GATTTATGGC	GGGCGTTGGT	CAGAGTACAT	ATTCAGGAAG	CTGCATTAT	540
	ACCTGGNAAA	CACCTGCTTT	ACTGTTGG				568

(2) INFORMATION FOR SEQ ID NO:178:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 573 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:178:

10

CTACNNTAAT	AAGAAAATAT	AACATACNAA	TCAAAAACATA	AAGGGATGTG	ACGTTAATGA	60
AACTCGTATT	TGTGGCACGA	GCTGGTAATA	TGGCACAAGC	TATATTTACA	GGAATTATTA	120
ACTCAAGCAA	CTTAGATGCC	AATGATATAT	ATTTAACAAA	TAAATCTAAT	GAACAAGCTT	180
TAAAAGCATT	CGCTGAAAAA	CTAGGTGTTA	ACTATAGTTA	TGATGATGCG	ACATTATTAA	240
15	AAGATGCAGA	TTATGTTTTT	TTAGGTACCA	AACCGCATGA	CTTTGATGCT	300
GCATCAAACC	ACATATCACA	AAAGACANTT	GCTTCATGTC	AATTATGGCA	GGTATTCCGA	360
CTGATTATAT	TAANCAACAA	TTAGAATGCC	AAAATCCAGN	TGCTAGAATT	ATGCCANACA	420
20	CAANTGCGCA	NGTTGGACAC	TCAGTTACTG	GCATTAGTTT	TTCAAACAAC	480
AAATCCTAAA	GATTAAATTA	ACGATTAGT	TAAAGCATT	GGTCTGTATT	GAAGTATCCA	540
GAGATCATTT	TACATCCAGT	TACAGCTATC	ACC			573

25

(2) INFORMATION FOR SEQ ID NO:179:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 619 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: Genomic DNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:179:

40

CTCTATTATA	AACATATTAA	AACGCATTTT	TCATGCCTAA	TTTATCTAAA	TATGCATTTT	60
GTAATTTTGT	AATATCACCT	GCACCCATAA	ATAAAACAAC	AGCATTATCA	AATTGTCTTA	120
ATACATTAAT	AAGAATCTTC	ATTAATGAAC	GATGCACCTC	CAATTTTATC	AATTAAATCT	180
TGTATCGTTA	ATGCGCCAGA	ATTTTCTCTA	ATTGAGCCAA	AAATTTTACA	TAAGAATACA	240
45	CGATCTGCTT	TACATAAACT	TTCTGCAAAT	TCATTTAAAA	ATGCTTGTGT	300
GTGTGTGGTT	GAAATACTGC	AACAACCTCT	TTATGTGGAT	ATTCTTTTCG	TGCTGTGTCA	360
ATTGTAGCAC	TAATTTCTCT	TGGATGGTGT	GCATAATCAT	CTACAATAAC	TTGATTGCA	420
ATTGTAGTTT	CATTGAAACG	ACGTTTAACA	CCACCAAACG	TTTCTAATGC	TTCTTTAATA	480
50	TTTGTAACAT	CTAGCTTCTC	TAAATAACTA	ATCGCAATTA	CAGCTAATGC	540
GTATGGTCAC	CATATTGTGG	AGACAGGAAG	TGATCATAAA	ACTCACCATC	CACATACACA	600
TCAAAAGCAG	TCTCGTGCC					619

55

(2) INFORMATION FOR SEQ ID NO:180:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 443 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:180:

```

GGCACGAGCG TGTCAATATT TACATGATGA TTTCATTAA TAGCCTTTTT TAGCATCAAC   60
TTCATTCTCA ATTAGACCAT TCTTATTGAG AAAATTAACT AGATTGTTTT TAAAAATATC   120
TAATAAGTCA TACTTTGCCT CATAATCATT ACCAGTTATA TGCGCTGTTA TAGTTACATT   180
TTCCAATTCA TATAATTCAT GATTAGGTTT CAAAGGTTCA TTTTCAAACA CATCTAAATA   240
TGCATGTCGA ATAACTTTAC TTTTAAATAC TTCTATTAAG AGCGCTTCCT TTAACATATG   300
TACCTCGTCC TATATTTATA AAAAGTGCTT CCATCCTTTC ATTAATTCCA AAATGTTTTT   360
TTCCTTTAGT TAAATGAATC CGTTTCCTTG CGTTTCCTGG TAAAGCATTT ATAATAATGT   420
CAGCATTTGG TAATGTGCTT TCA                                     443

```

## (2) INFORMATION FOR SEQ ID NO:181:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 569 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:181:

```

GGCACGAGCT ATGTGGTTTG AAGTCAATCG CCAAATGTTC CACCTACCAT TCTGATAAAA   60
GTATTTACTT GTTCTTTTAT CCATAACACA TATACTTTAT GATCTCTATT TTCAAATTGT   120
CTATGCACAT ATTTTTGTAA AGGATGCAAC TTTCCTTTT CTTGCTTCAT TTCTACAAAA   180
TATGTTTTTC CTTCTGGCAT AATAATAATT CTATCTGGCA CACCTCTTGT TCCAGGTGCG   240
ACCCATTTTA AACATAAACC GTTTAGCTTT GTTATCTCTT TCACTAAATA TTTTCTAAT   300
GTCGATTCCT TTCATTTTAT CACCTTGAT ACAAATTTA TATTGTGTT CCGAATGTTT   360
GTTATCAATT CCTTGCCAAA CTTTTTAAAA ATAGCTGTTT AGAGGGTTTA CCCCTATACC   420
CCCTTTACTC CCCTAAACAC TACTTTTTTA AACTTTATAG TGAATTTGAA TGCAACATTG   480
GGAAACAAAC AGGTTTGAAC CCCTACAGCT AGAAAGAAGA GTGTTTGTA TCATTGTTGC   540
ATCATGTTGC ATCACC AAAA TGATACAAC                                     569

```

## (2) INFORMATION FOR SEQ ID NO:182:

# EP 0 841 394 A2

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 511 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:182:

GGCACGAGCG	CAAAATATTT	CAAATAAAAA	TGATTGTAAA	AAGGCAAAAT	ACAAAATTTT	60
ACTTAACAAC	TAGTACATAA	AGTAATACAA	TTAAATTAAT	TCTATCTGAA	AGATGTGTGG	120
GGCATCGTTA	TTTTAGGTGG	ATATGAGCAA	TTTATTAAAA	GTCATTTACG	GAAAATATAT	180
ATAGACGGGG	TGAGTAATAT	GCAAGAACAT	GTGGTGGTTA	CACTTGATGG	AAAAGATTAT	240
CTTGTAGAAC	CAGGTACGAA	TTTACTTGAA	TTTATTAAAT	CACAAGATAC	TTTGTCCCT	300
TCAATTTGTT	ATAACGAGTC	GATGGGCCCA	ATTCAAACAT	GTGATACATG	TACTGTTGAG	360
ATTGACGGTA	AAATTGAACG	CTCATGTAGT	ACGGTGATTG	ATCGCCCAAT	GACTGTAAAT	420
ACTGTGAACA	ATGATGTGAA	AGATGCTCAA	AAAGAGCCTT	GATCCGAATT	TTAGAAAAGC	480
ATATGCTGTA	TTGGACAGTA	TGTGATTATT	A			511

## (2) INFORMATION FOR SEQ ID NO:183:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1125 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:183:

ATCCTTTATT	TNTAAGGCGT	TTCATCAAGC	TAACACTTCA	TTATCTTAG	TCGCTTTAAT	60
GCCCTTCTAT	TAAACTCGGA	ACTAAAAATG	GATTGNCTTT	TCAACCACCC	GANCACTAAA	120
ACATTGTCNT	TTTTGATGGC	CNATTAAGAC	ACAATTTCCC	GCTGTTTGCT	TCAAAGTAGC	180
TTGCTTCTTG	NATTNATTTT	CAATATCTTT	CTTGTTAAAA	ACAAGANTGT	TGCACAGTTT	240
GATTGGCATC	TTTATTTAGA	NCAATGGCAT	CTGCTTGCCA	CTTATCAATG	CCTTCTTTAT	300
TCATATTGAT	AAGACCATTC	GCCAATCCAG	ATAATAAAAA	TAGCAAGTAA	CTAATCATCG	360
TTAACACACC	AATAATTAGT	CCAAACTTCA	ATTTGTTGCG	CCGTATTTCA	TTCCAAGCTA	420
AAAACATGCA	TTTCTCTCCC	TACTACTATG	ATTTAAACAT	TGTTTATATT	CTTAGATGCA	480
CGTACGTCGT	GTTGCGCTCT	GTAATGTTAT	ACATACACTT	ATCCTTCATT	ATACCCGANC	540
TTTTTATATT	AAAACCAAAT	TTATGGAAAA	TGCAANNANT	TGTCATTATT	TTTTGTGCGG	600
TACATTTAAA	ATTAAGGATC	AATTTAAAAA	CGCTTACATA	TACCTTTAAG	TACATGAAGA	660
CGTCCAATTC	ATATATTATT	TAACTTCGCC	TGTTTTAGGA	TCGGATTGCT	TAATAGCATT	720



EP 0 841 394 A2

	TTTACGTAAT	TTATCTTTTG	CTTTGTCAC	TGCTTTATAG	TTATTGTTGT	AAATCGTAGC	780
	TTCCCAACTA	CCATACATTG	GGTTAGGGAA	AATGATATAT	TTCTTACCGA	AATCGTCTTT	840
5	ATGTTTTTCA	ATTAATGCTT	CACGAGATTC	AGCTGTAGCT	TCTTTTGGAT	CTGTAAAGTC	900
	TAATAAATTA	TCTCCAAATA	GCATGACAAG	TTTATGATCC	TTTTGAACCA	TTTGTCTGCG	960
	TGATTCTTTA	CTCTTATCAT	CTTTACCTTT	TAGTAAAATA	TGACTCTTCT	TAGCTTGAGG	1020
	GATACCTTGT	TGTTTTAAGT	TCTTTTGTGT	TGCCTTTAAA	TCTTTTCTTT	TATCTCTATC	1080
10	AGAAATATAG	TAGATATCGA	CACCTTTTTT	GTCAGCATAT	TTCAA		1125

(2) INFORMATION FOR SEQ ID NO:184:

(i) SEQUENCE CHARACTERISTICS:

- |    |                             |
|----|-----------------------------|
| 15 | (A) LENGTH: 3100 base pairs |
|    | (B) TYPE: nucleic acid      |
|    | (C) STRANDEDNESS: single    |
| 20 | (D) TOPOLOGY: linear        |

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:184:

	AGAAGCATAT	AAACAAAACA	TGTTGCGATT	ATTAGGTAAA	ACTGGTTTTG	AAGACTTGAA	60
	AAAAGAATTA	GAAGAGCGTT	TATAAAATAC	ATTACTTCAA	TGATTAGTGA	AGTTTGAAAA	120
30	GATAGAACTA	GACGTTAACT	ATTTAAAGCA	TATTTTCGAG	GTTGTCAATTA	CAAAATGTA	180
	AATGTAATGA	CAACCTCGTT	TTTATTTATA	TGCAAGAACT	AGGTTACTAG	CTAATGTGAC	240
	AAGATGTTAA	GAGAAAATTA	AAGAAAAAAT	AACATCTGTC	ATACAATAAT	ATTGTTATAC	300
	TACTAGAGAC	TGATTTATTA	GCATGATTAC	ATGTTAATGT	TTCTTTACTT	AGTAATTAAC	360
35	TTTATAATGT	AAGAATAATT	ATCTTCAACC	AAAGAAAGGG	ATTGATGATT	TGTCGTTTCA	420
	TCAAGTAGAA	GAATGGTTTG	AGATATTTTCG	ACAGTTTGGT	TATTTACCTG	GATTTATATT	480
	GTTATATATT	AGAGCGATAA	TTCCAGTATT	TCCTTTAGCA	CTCTATATTT	TAATTAACAT	540
	TCAAGCTTAT	GGACCTATTT	TAGGTATATT	GATTAGTTGG	CTTGGATTAA	TTTC TGGAAC	600
40	ATTTACAGTC	TATTTGATCT	TGTAAACGAT	TGGTGAACAC	TGAGAGGATG	CAGCGAATTA	660
	AACAACGTAC	TGCTGTTCAA	ACGCTTGATT	AGTTTATTATG	ATCGCCAAGG	ATTAATCCCA	720
	TTGTTTATTT	TACTTGGGNT	TTCCTTTTC	GCCAAATACA	TAAATAAATT	TTGTAGCGAG	780
	TCTATCTCAT	ATTAGACCTA	AATATTATTT	CATTGTTTTG	GCATCATCAA	AGTTAGTTTC	840
45	AACAATTATT	TTAGGTTATT	TAGGTAAGGA	AATTACTACA	ATTTTAAACG	ATCCTTTAAG	900
	AGGGATATTA	ATGTTAGTTG	TGTTGGTTGT	ATTTTGGATT	GTTGGAAAAA	AGTTAGAACA	960
	GCATTTTATG	GGATCGAAAA	AGGAGTGACA	TCGTGAAAAA	AGTTGTAAAA	TATTTGATTT	1020
	CATTGATACT	TGCTATTATC	ATTGTACTGT	TCGTACAAAC	TTTTGTAAATA	GTTGGTCATG	1080
50	TCATTCCGAA	TAATGATATG	TCGCCAACCC	TTAACAAAGG	GGATCGTGTT	ATTGTAAATA	1140
	AAATTAAAGT	TACATTTAAT	CAATTGAATA	ATGGTGATAT	CATTACATAT	AGGCGTGGTA	1200
	ACGAGATATA	TACTAGTCGA	ATTATTGCCA	AACCTGGTCA	ATCAATGGCG	TTTCGTGAGG	1260
	GACAATTATA	CCGTGATGAC	CGACCGGTTG	ACGCATCTTA	TGCCAAGAAC	AGAAAAATTA	1320
55	AAGATTTTAG	TTTGCGCAAT	TTTAAAGAAT	TAGATGGAGA	TATTATACCG	CCTAACAATT	1380
	TTGTTGTGCT	AAATGATCAT	GATAACAATC	AGCATGATTG	TAGACAATTT	GGTTTAATTG	1440

EP 0 841 394 A2

	ATAAAAAGGA	TATTATTGGT	AATATAAGTT	TGAGATATTA	TCCTTTTTCA	AAATGGACGA	1500
	TTCAGTTCAA	ATCTTAAAAA	GAGGTGTCAA	AATTGAAAAA	AGAATTATTG	GAATGGATTA	1560
5	TTTCAATTGC	AGTCGCTTTT	GTCAATTTAT	TTATAGTAGG	TAAATTTATT	GTTACACCAT	1620
	ATACAATTAA	AGGTGAATCA	ATGGATCCAA	CTTTGAAAGA	TGGCGAGCGA	GTAGCTGTAA	1680
	ACATTATTGG	ATATAAAACA	GGTGGTTTGG	AAAAAGGTAA	TGTAGTTGTC	TTCCATGCAA	1740
	ACAAAAATGA	TGACTATGTT	AAACGTGTCA	TCGGTGTTC	TGGTGATAAA	GTAGAATATA	1800
10	AAAAATGATAC	ATTATATGTC	AATGGTAAAA	AACAAGATGA	ACCATATTTA	AACTATAATT	1860
	TAAACATAA	ACAAGGTGAT	TACATTACTG	GGACTTTCCA	AGTTAAAGAT	TTACCCGAAT	1920
	GCGAATCCCA	AATCAAATGT	CAATCCAAAA	GGGTAAATAT	TTAGCTCTTG	GAGGATAATC	1980
	GTGAAGTAAG	TAAAGATAGC	CGTGCGTTTG	GCCTCATTGA	TGAAGACCAA	ATTGTTGGTA	2040
15	AAGTTTCATT	TAGATTCTGG	CCATTTAGTG	AATTTAAACA	TAATTTCAAT	CCTGAAAAATA	2100
	CTAAAAATTA	ATATGAAACA	AATACAACAT	CGTTTGTCTG	TTTTAATACT	GATAAACGAT	2160
	GTTTTATTTG	GTTAGTACCA	CAATAAAAGC	TAAGTTCGAA	ATGAACCTAT	AATAAATCAA	2220
	TCACAATCAC	TTTGTGTAA	AATATGTGTC	AAAGGAAGTG	AGGGTTTGTC	ATGACATTAC	2280
20	ATGCTTATTT	AGGTAGAGCG	GGAACAGGTA	AGTCTACGAA	AATGTTGACC	GAAATAAAAC	2340
	AAAAAATGAA	AGCAGATCCG	CTTGAGATC	CAATCATTTT	AATTGCGCCA	ACTCAAAGTA	2400
	CATTTCAAAT	AGAACAAGCC	TTTGTCAATG	ATCCGGAATT	AAATGGTAGT	TTAAGAACAG	2460
	AAGTGTGCA	TTTGAACGA	TTAAGTCATC	GTATTTTCCA	AGAAGTTGGT	AGTTATAGCG	2520
25	AACAAAAGTT	ATCTAAAGCT	GCAACGGA	TGATGATTTA	TAACATTGTT	CAAGAACAAC	2580
	AAAAGTATTT	AAAACTTTAT	CAATCACAAG	CAAAATATTA	TGGGTTTAGT	GAAAAATTAA	2640
	CAGAACAAAT	TCAAGATTTT	AAAAAATATG	CAGTAACGCC	TGAACATTTA	GAACACTTTA	2700
	TTGCTGATAA	AAATATGCAA	ACTCGAACTA	AAAATAAGTT	AGAGGATATT	GCTTTAATAT	2760
30	ACCGTGAGTT	CGAACACGC	ATNCCANAAC	GAGTTTATTA	CTGNTGAGGA	TTCATNACAA	2820
	TATTTTATTG	ATTGTATGCC	GAAATCAGAG	TGGCTAAAAC	GTGCTGATAG	ATATATTGAT	2880
	GGTTTTCACA	ACTTTTCAAC	GATTGAGTAT	TTAATAATCA	AAGGATTAAT	TAAATATGCG	2940
	AAGAGTGTCA	CAATTATATT	GACGACAGAT	GGTAACCACG	ATCAATTTAG	TTTATTTAGA	3000
35	AAACCATCGG	AAGTGTTACG	ACATATTGAA	GAAATAGCAA	ATGAACCTAA	TATTTCTATT	3060
	GAACGTCAAT	ATTTCACCA	ATTATATCGC	TTCAATAATC			3100

(2) INFORMATION FOR SEQ ID NO:185:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 817 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:185:

	GGTGGCGAGA	AAATACCGCA	AGGTCATAAA	GATATCTTTG	ATCCAACTT	ACCAACAGAT	60
55	CAAACGGAAA	AAGTACCAGG	TAAACCAGGA	ATCAAGAATC	CAGACACAGG	AAAAGTGATC	120
	GAAGAGCCAG	TGGATGATGT	GATTAAACAC	GGACCAAAA	CGGGTACACC	AGAAACAAAA	180

# EP 0 841 394 A2

	ACAGTAGAGA	TACCGTNTGA	AACAAAACGT	GAGTTTAATC	CAAAAATTACA	ACCTGGTGAA	240
	GAGCGAGTGA	AACAAGAAAG	ACAACCAGGA	AGTAAGACAA	TCACAACACC	AATCACAGTG	300
5	AACCCATTAA	CAGGTGAAAA	AGTTGGCGAG	GGTCACCCAA	CAGAAGAGAT	CACAAAACAN	360
	CCAGTAGATA	AGATTGTAGA	GTTCGGGTGG	AGAGAAACCA	AAAGGTCCCA	AANGGACCTG	420
	AAAACCCAGA	GAAGCCGAGC	AGACCAACTC	ATCCANGTGG	GCCAGTAAAT	CCTAACAAATC	480
	CAGGATTATC	GANAGACAGA	GCAAAACCAA	ATGGCCCCAGG	TCCATTCAAT	TGGATAAAAA	540
10	TGATAAAGGT	TAAAAAATCT	AAAATTGCTA	AAGAATCAGT	AGCTAATCAA	GAGAAAAAAC	600
	GAGCAGAATT	ACCAAAAAACA	GGTTTAGAAA	GCACGCAAAA	AGGTTTGATC	TTTAGTAGTA	660
	TAATTGGAAT	TGCTGGATTA	ATGTTATTGG	CTCGTAGAAG	AAAGAATTAA	AATAATTCAT	720
	AATTTAAATA	ATAGTTGATT	TGCATTCACT	ATATTTAGTT	TGTTAAAAAC	AACCTAGAAT	780
15	ATGATGAGAA	TGATATACAA	CCCCAAAAGT	TGGCTTG			817

(2) INFORMATION FOR SEQ ID NO:186:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1348 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:186:

	CCATCGACTA	AGCAATGGTA	TTTGTTTCATT	AAATCCTCAT	GTTGATCGTT	TAACGTTAAG	60
	CTGTGCGATG	GAAATCGATG	CTAGTGGTCG	CGTTGTTAAA	CATGAAATTT	NTGATAGTGT	120
35	TATACATTCT	GATTATCGAA	TGACGTATGA	TGCGGTAAAT	CAGATTATTA	CTGAAAAGGA	180
	TCCTAACATT	CGCGAACAA	ATAAAGAAAT	TACGCCTATG	TTAGATTTAG	CACAAGATTT	240
	ATCTAATCGN	TTGATTCAAA	TGAGAAAACG	ACGTGGGGTG	AAATCGATTT	TGGATATTAG	300
	NGAAGCAAAA	GTATTAGTTA	ACGAAGTCGG	GTATACCAAC	AGATGTTCAA	TTAAGACAAC	360
40	GTGGCGAGGG	TGAACGTCTA	ATTGAATCAT	TTATGTTAAT	TGCAAAATGAA	ACAGTTGCTG	420
	AACATTTTAG	TAAGTTAAAT	GTACCTTTTA	TTTACCGAGT	GCATGAGCAA	CCTAAATCAG	480
	ATCGCTTAAG	ACAATTCCTT	GATTTTATTA	CAAAC'TTGG	CATCATGATT	AAGGGCACTG	540
	GCGAAGATAT	TCATCCAACA	ACACTTCAAA	AGGTTCAAGA	AGAAGTAGAA	GGTCGACCTG	600
45	AACAAATGGT	CATTTCAACA	ATGATGTTAC	GTTCAATGCA	ACAAGCGCAT	TATGATGATG	660
	TGAACCTGGG	ACATTGTGGC	TTATCAGCTG	AATATTATAC	GCATTTNACA	TCACCAATTA	720
	GACGTTATCC	TGATTTAACA	GNTCATCGTT	TAATCCGTAA	GTATTTAATT	GAGAAATCAA	780
	TGGATAACAA	AGAAGTGAAG	CGTTGGGAAG	ACAAATTGCC	TGAGTTAGCT	GAACATACTT	840
50	CTAAACGTGA	ACGTCGTGCT	ATTGAGGCAG	AACGTGATAC	TGATGAATTG	AAAAAAGCAG	900
	AATATATGAT	TCAACATATT	GGTGATGAAT	TTGAAGGTAT	TGTCAGCTCA	GTAGCTAACT	960
	TCGGTATNTT	CATTGAATTG	NCAAATACGA	TAGAAGGTAT	GGNTCATATT	GCGAATATGA	1020
	CTGATGATTA	TTACCGCGTT	GAAGAGCGTC	AAATGGCATT	AATTGGGTGA	GCGTCAAGCT	1080
55	AAAGTATTTA	GAATTGGTGA	CACAGTTAAG	GTAAAGTGA	CGCATGTTGA	TGTAGATGAA	1140
	CGATTAATTG	ATTTTCAAAT	TGTTGGAATG	CCTTTACCTA	AAAATGACCG	CTCACAGCGA	1200

EP 0 841 394 A2

5      CCAGCAAGAG GTAAAACGAT TCAAGCTAAA ACGCGTGGCA AATCTTTAGA TAAATCGAAA 1260  
TCTGATGATA AGGGTCGGTA AGAAAAAAGG TAAGCAACGT AAGGTAAAAA CCAACGTAAT 1320  
AATGATAATC AGGTAATAGT AAGCATAG 1348

(2) INFORMATION FOR SEQ ID NO:187:

10      (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 982 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
15      (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

20      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:187:

ATGAAATTA TTTTAATTTT AGCATTAGCG CGTGTCGTTT CTAGACATAA TCAATTCACA 60  
TTCAATAAAT CATTCCAAAG TGATTTGTTA TTATTTTTC AATTTATTGG TGTCTCGTTA 120  
25 GTACCAAGTA TTTTAATATT ACTGCAAAAT GACCTAGGAA CTACATTAGT ATTAGCTGCT 180  
ATTATTGCAG GTGTGATGTT AGTAAGTGGT ATAACATGGC GTATCTTAGC ACCTATCTTT 240  
ATTACAGGTA TTGTTGGTGC AATGACAGTC ATTTTAGGTA TTCTATATGC ACCCGCATT 300  
ATTGAAAATT TATTAGGTGT CCAACTGTAT CAAATGGGAC GAATCAATTC AAGGCTTGAC 360  
30 CCCTATACAT ATAGTAGGGG GGGGGATGGC TATCAATTAA CTGAATCACT TAAAGCTATC 420  
GGGCTCTGGA CAAGGTACTA GGTAAAGGAT ACAATCACGG GGGAGGTTTA TATCCCTGA 480  
AAATCATACT GACTTTTATC TTTNCAAGNG AATGGGAGAG GAACCTGGCT TTATCGGTTC 540  
TGGCAAATTG AGNCTTAATA TTTTATTTT TAATCTTCCA TCTAATAAGA TTAGCTGCGA 600  
35 AAATTGGAGA TCAATTTACC AAAATCTTTA TCGTTGGTTT CGTCACTTTA CTTGTGTTCC 660  
ATATTTTACA AAATATTGGT ATGACAATTC AGTTGTTACC AATCACTGGT ATTCCATTAC 720  
CATTTATTAG TTATGGTGGT AGTGGCTAT GGAGTATGAT GACTGGAATA GGTATAGTCT 780  
TATCAATTTA TTATCATGAA CCAAACGAT ATGTCGATTT ATACCATCCA AAAAGTAATT 840  
40 AATTTAAACT ATTTTGAGTT TCAAATATCA TAACTTTTCA AGATGACGTT ATATAGTCTA 900  
TTTACGTCGT CGATTTAAAA TGTCATATAT AGATATTACT CGATAATAAC AATCCCTCTT 960  
TGAAGTACAC ATTGTAAAAT GG 982

45      (2) INFORMATION FOR SEQ ID NO:188:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1709 base pairs  
50      (B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

55      (ii) MOLECULE TYPE: Genomic DNA

EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:188:

5	GGCACTAACT TCATTTGAGC ATCCTCCTAT CAATTGCTAT ATAAATTAGT ACCCTTTTGC	60
	CACTTAATTA TAACAAATTC TCAAATTTTA AAAATTGAAA ATCTAGTTAA TGTATTAGCT	120
	CGATTTTGAA ATCTAATAAT AATTGGCATA AAATGGAAGT AATATTATGT GAGGAGTGTT	180
	ATGAAATGAC AAAAATATCA AAAATAATAG ACGAACTGAA CAATCAACAA GCTGATGCAG	240
10	CATGGATTAC AACACCGTTG AATGTATATT ATTTTACTGG ATACCGTAGC GAACCCATGA	300
	AAGATTATTT GCATTATTGA TTAAGAAAGA TGGTAAACAA GGTACTATTT TGTCCAAANA	360
	TTGGAAGTCG GAAGAAATCA AAGCATCACC CTTACAGGTT TGAAATCGTT GGATATTTAA	420
	GACACTTGAA ATCCTTTTTC ACTTTATCCA CAACNCAATC AATAAATTAC CTAATTGTAA	480
15	GCGGAGCACC TTAACAAAGTA GCACCGCCAC AAACAATTAA TCTCTGTTTC NATGTCAATT	540
	CATTTCGAGA TGTGAGTTA ACAATCACAC AATTAAGAGA TATTAAATCC GAAGATGAAA	600
	TTAGCACAAAT ACGTAAAGCT GCTGAGTTAG CAGATAAGTG TATCGAAATA GGTGTTTCTT	660
	ATTTAAAAGA AGGTGTGACT GNACGTGAAG TAGTCAACCA TATTGAGCAA ACTATCANAC	720
20	AATATGGCGT CAATGAAATG AGTTTTGATA CGATGGTTTT ATTTGGAGAT CATNCCGCAT	780
	CACCTCATGG CACACCAGGA GATCGCAGAT TAAAAAGCAA TGAATATGTA CTATTTGATT	840
	TAGGTGTAAT TTATGAGCAT TATTGTAGCG ATATGACACG TACCATTAAA TTTTGGTGAA	900
	CCTAGCAAAG AAGCACNAGA AATTTATAAT ATTGTATTAG AAGCAGAAAC ATCTGCAATC	960
25	CAAGCAATTA AACCTGGAAT ACCTTTAAAA GATATCGATC ATATCGCTAG AAATATTATT	1020
	TCAGANNAAG GTTATGGTGA ATATTTCCCT CATCGCTTAG GTCATGGTTT AGGATTACAA	1080
	GAACATGAAT ATCAAGATGT TTCAAGTACT AATTCTAATT TGTTAGAAGC TGGCATGGTT	1140
	ATTACAATCG AACCAGGTAT TTATGTACCA GGTGTTGCAG GTGTAAGAAT TGAAGATGAC	1200
30	ATACTTGTC CTAATGAAGG ATATGAAGTA TTAACACATT ACGAAAAATA AGGAGTGGGA	1260
	TAAAAATGAA AAGCTTGTTA CAAGCACATT CTCATTCACT CAAACACTGC CAATATAACA	1320
	TTGTAGCGCC TAAGACATAA ATTTTATCC AAGTCTAAAT GCAATATGTA ACAAACAAGC	1380
	TAGAAACACA TATGCAGGTA TGTTCATCAG TAACATGTAA TGAATCAAAT CAATATCATT	1440
35	CATGTTTCGAT GATTTCCTCG CATGTGTTCT AGCTTTAATT TATCATTATT TAATTTTAAT	1500
	AACCAAGGAG ATGATAACGT CAGTCTTTAG TACGCTGTAA TCCATTCCCT TTTCATCAAA	1560
	TTCAAATTAT AATTGTAATG CTTCTTCTAC AGATTTATAT TCCATTTCAA ATGCCTCTGC	1620
	AACGCCTTTA TTGGTTACGT GACCTTTGTA AGTATTTAAA CCTAATGATA ATGGTTGATT	1680
40	TGATTTAAAT GCTTCTCTAT ACCCTTTAG	1709

(2) INFORMATION FOR SEQ ID NO:189:

45 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1355 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

50 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:189:

EP 0 841 394 A2

	AAAGAGGAAG	ATGTGAACCA	CCACTAGTAT	TGAATCTAAA	GATTGAGGAC	GTCCTTTAT	60
	TAATTGGCTA	NNNAAGNNAN	ATCNNAAGA	TAGCCAATCC	AAAGCCAAAA	AGTCAGCATC	120
5	AAGAAATACT	TCTANNAAGG	TAGCAGCTNN	NNAGAAGNGN	NAGAAATCTA	AGANAAATNN	180
	NAAATAATTT	GTTTCTTTGC	TAAATAGAGG	AGCACCGATT	GACATCACAT	CAGTCGGTGC	240
	TCCTTTTATT	TATCTTTTTT	AATTAATTTA	TACAAANCCN	ATTCCCTGTT	GAGCGTGTG	300
	AATCGCCTTC	CNTTGTTTTG	TTCTCNCGGT	ANCCCATTTA	ACCAAATNAT	AAACCNANAT	360
10	CTTTNTCCAA	ATATTTCTAT	TTGATCAAAA	TAAGGTTTGA	AATTTGCGTT	TTTCACATAA	420
	CCAGCTCGTG	CCAATGCTAT	CGTGCAATTA	GCTTTGAGTC	TGTATATAAT	AGTGCGTTTT	480
	GAACATTTAA	TTACAGTGCA	TGTTCTAGTG	CATAAATACA	TGCAGCCCAT	TCTGCAGTGT	540
	GGTTATCCAT	TTCGCCTAAC	TCATGTGTAT	ATGTATNATG	CTGCTTATCT	TCTTTGATTA	600
15	CAATGGCACA	NGTACTTATG	CCTGGATTTT	CTNNTCGTCG	CAGCATCAAA	ATTTATGTGC	660
	GCCATAATAA	ACCTACTTTC	TATTCAATAC	TTAGTTAAAG	TTACTATTAC	TGTAATACAA	720
	AATATGTTGG	GTAATCCATT	AAAAACACG	CATCACTTAA	ATAAGTAAAC	ACGTTGTTAA	780
	AAATACTTCG	CTTGATTCAA	AAGATGATTT	TCTAAATACG	TAGTNCCTGT	AAAATACTTC	840
20	CTAAANAAAT	CATCTTCAGG	CTGGGACAT	AAATCAATGT	TCTATGCTCC	TNCCGAAGTT	900
	ATATTGGCAG	TAGTTGACTG	ANCGAAAATG	CGCTTGTAAC	AAGCTTTTTT	CAATTCTAGT	960
	CAGGGGCCCC	AACACAGAAG	CTGNCGAAAA	GTCAGCTGAC	AATAATGTGC	AAGTTGGGGA	1020
	TGGACCCCAN	CAAAGAGAAA	TTGTATTCCC	AAATTCTACA	GACNATNCAA	GTGGGGTGG	1080
25	GNCGACGAAA	TAAATTTTGC	GAAAAATCA	TTTCTGTCCC	ACTCCCTTAA	AACTTATTCT	1140
	TTTGTGTAGT	AAGTGCGTTA	ATAGCCTTGA	TCTAACTTAT	CAATCTTACC	TTTACGATAA	1200
	AATGATTTAG	CAATATATCC	ANNTGGTACA	TTGAAAAC TG	TTGAAGCTAA	TTTTAATACG	1260
	TAAGTTGTAA	TAAATATTTT	NAATNCAACT	GTACCAGGTA	AACTCCGAT	AAAAGCGATA	1320
30	GCTACAAATA	AAGCTGTATC	CAATTATTGA	GCTTG			1355

(2) INFORMATION FOR SEQ ID NO:190:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1410 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:190:

	CCAATTTTCC	CTTCTACCTG	CGTGGGTAGT	ATTCAAGCCC	AAAGTTCAGC	GGGAAAATAT	60
	TGACAATTAA	CCGGATCCAG	CTAATCGGAG	GTCGGATAAA	GATGTTGAGT	TGTAAGTTAT	120
50	TTGGAATATT	ATTTTAATAG	TGTCATCCCC	CTTTGTAAAA	TAATTGTCTT	ACTTTTAAAT	180
	TAAAAGCCAA	ATTAATATAA	GAAANCTAAG	ACTTAGTACN	GTATCAATTT	TGTGCGTTTC	240
	AATTGAGTTC	TAGTTTTTTT	TAATATGTTA	ATATTAAACT	TATAACTTTA	TGGGAGTGGG	300
	ACAAGAATGA	TAAAGAGCCA	CTAAATGATT	TATTATGTAG	TGGTTCTTAA	TACATTAGCC	360
55	ACANCTAATG	TGTACTTAAA	AATAGGAATA	CATGAGTAAA	ACTCATGCAT	AAGAAATACT	420
	AATTTCTATA	GAAAAAGTAT	TACTTTATCG	TTGTACCACA	CCAACCTGCA	CATTATCGTA	480

EP 0 841 394 A2

5 AGCTGACTTA TCGTAAGCTT CTGTGTTGGG GCCACACCC CAACTCGCAT TGCCTGTAGA 540  
 ATTTCTTTTC GAAATTCTCT TTGTTGGGGC CCACACCCCA ACTTGCATTG TCTGAAGAAA 600  
 TTGGAAATCC AATTTNCTCT GTGTTGGGGC CCACACCCCA ACTCGCATTG CCTGGAAGCT 660  
 GAATTTCTTT TCGAAATTCA GCTTCTGTGT TGGGGCCAC ACCCCAACCT GCATTGCCTG 720  
 TAGAAATTC TTTGAAATC CAATTTCTCT GTGTTGGGGC CCCTGACTAG GATTGAAAAA 780  
 AGCTTGTTAC AAGCGCATTT TCGTTCAGTC AACTACTGCC AATATAACTT CGTAGAGCAT 840  
 10 AGAACATTGA TTTATGTCCC AGCCTGATAT CACCATTAAA TACAATTCAT TTAGTNTTCA 900  
 ATTGGAACA ATTNATCGAT ATATTGAATC TCATCATCTG ATAAAACGAT ATCTGCAGCT 960  
 TTAATATTTT CAATGACTTG TTCTGCACGT TTTGCCACAG GAATAATCAC ATCGATAGCT 1020  
 GGTCTCGTTA AATAAAATGC TAATACAATG TTCGCAATTG AAGTTTGATG TGCTGCAGNT 1080  
 15 AGNCTTTTCC AAAGCTTTTA CGCTCGTGAC GCACATTCTC TTCAAAAACA CCAGGTATAA 1140  
 AATCCCGAGC TGTTAGTACGA TGGTCACTAA ATTTAGTGTT CTCATCATAT TTTCAGCTA 1200  
 AAATACCGGA TGCTAATGGG AAATAAGGAA TAAATGTGAT TAGGTGATCA ACACAATATT 1260  
 GCAATACTGC CTCATTTTCG CGGTGCAATA AATTATATTC TAACTGTACA ACATCAACGT 1320  
 20 AACCATCTTT ATTTGCTTCT TTAAGTTGAT CTAATGTGAA ATTTGATACA CCAATGGCTT 1380  
 TAATCTTCCC TTGTTCCCTA AGCTCGTGCC 1410

(2) INFORMATION FOR SEQ ID NO:191:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2437 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:191:

40 GTGTTACGGT AGGTGTTGA TCCTAATTTT CCCTGNCAGT TGGTCCCAGT ACCCACCCAT 60  
 TTCCCAACTT ACTAAAAAAG AAATCCAATT ANNCCGAAAG AAAATATTGA AGTCATTGAA 120  
 ATTAACCGAA GCGTTCCAGT GCACAGGTAG TTGCCTCCCC AACCAAGCTT TAAATATTTT 180  
 CAAATACGCA ATTAAATATA TGGGGTGGTG CATTAGCATC AGGTCATCCA TACGGTGCAA 240  
 GCGGTGCCCA ATTAGTGACT CGATTATTTT ATATGTTTGA CAAAGAGACT ATGATTGCAT 300  
 45 CTATGGGGAT AGGGGAGGT CTAGGAAATG CAGCATTATT TACTCGATTC TAACCAGCGA 360  
 TTAAATGTGT CATTTCTTAA GGATAGTGTG GCTGCATATT ATCAGTGTTC TANCCAACCT 420  
 TATAGAAAAG AAGTCTCGTG CCCATTAAATG TGTNCGTCAT TATGGCCACA ATTTGATTTA 480  
 TTTAAAAAAA TATCCAAATA GCGAGCTGAT TTTACCACAA ATCAGCAATT AATCAAATC 540  
 50 NCAAGATAGA AGTAGACACA ATATATGTAG GGCATTTAGA AGATATTGAA TGCCGACAGA 600  
 CTCGCAATAT CACACGTTAT ACAATGGCTT TAACATTAAC TAAAAATGAT CAACATGTCA 660  
 TANC GTTAC ACAAACTTTT ATTAAGGCGA TGAAGTAGAG ATGAAGTTTA ATGAGATATG 720  
 GATAAATGAA TATTGGCGC TCGTAAATGA TGATAATCCA ATACATAATG AGATTGTGCC 780  
 55 AGGACAATTA GTGAGTCAAA TGATGCTGAT GGCTATGTCA TTAGAGACAA ACCAGTGTCA 840  
 AATTAAC TAC GTTAAACCTA TTTTAATAAA TGAAAATATC GAATTCATTG AACAAACGCA 900

# EP 0 841 394 A2

```

ACACGAAATT ATAGCAATTA ATGACGATGG AGAGATTAAA ATAAAAATTT CTTTGAGCAC 960
AAAAAATAA CCGATATTAG CTGCATGAAC GCATATTAAT TAGGAGATGA AAGGACAGCT 1020
5 AATATCAGTT ATGTATTGTT ATTATTATTG GGAACAGAGA TGAATATAGG TTACGTTTCT 1080
TTCTTTGCAC GGGGATGCAT TAATCTAAAA TAATAATAAC AACTATATCA ATGTTTAATA 1140
AATCTGGAT TATTGGAACG ATTAGTCAAT TTAACTAAC TNCATATGAT CTATATCGTC 1200
TTGTNATAAA GAGAGCAATT TGAATATTTC AGTATCACTA AATGAATCGN CACATTTAAT 1260
10 TGAAACATGC TGAAACGTTT GGGTTATAAT TTCATAAACT GGTGCGCCTT CATGGTGATA 1320
CTGTCGAATA AATAATCATA ACCATATTTA CCTCCTTNGG CTACTIONATG GGTATATTAT 1380
AAATAACATT TTTATGTGTG ACATCAACCT TAAGTATCAA CTTTTTATCA GACATAGAAC 1440
GTANGATTTA CTAAGACTAT TTATGTATAA AAGTTCTAAA TAAATATATA TTTATAGAGT 1500
15 CGCCTGGCAG NCATTTGGGA AATATAACAT ATATGATTAG AGAGGCATCT ATCGCAAAAG 1560
AATGATAATG ATAGAGGTAT TGAGCATATA GATGAGTTTA AGTTCATCTT GAAAATAAAG 1620
GGTTATTTAG TCATAGATGT AGATGTATAG GAAATATTTG TATGTATTGN TCGATATGTA 1680
TGAAATTTTC AATAAAAGCT AATAACGCTT ATATGTAACT TTCAAATTTA AATTATATAC 1740
20 AGAGCATGAT GATTATAAAA AAATANCCAC ATCACATAAA TTGAGTTCAT ACCCAATTTA 1800
AGTGGTGTGG CTAATAATGT TGATTTATAG ATGAACCGCC TAATCGTTAA ACCTCTGTTA 1860
CTTCAACATC GATATGTTCA ATACGGTTGT ATGCACCGTG ATCCACAGGA CCAACAAAAT 1920
CATTCAATTT CCAACCGTTT TTAATAGCAG AAGCGACGAA AGCTTTGCTT CGTGCTAATC 1980
25 ACAGCTTCTT TCGGTGACTT ACCGTTAGCT AAATATGCAG GTGTTGCCGC AGCAAATGTA 2040
CAACCAGCAC CATGGTTATA ACTTTGTTGG AACATGTCTG TTGTTAGTTG ATAAAATGTG 2100
TGACCATCAT AGTATAAGTC ATACGATTTA TCTTGATCTA AAGCTNTGNC ACCTTTAATG 2160
ATGACATGCT GTGCGCCTTT ATCAAAGATA ATTGNTGCAG CCTTTNACAT ATCTTCAATT 2220
30 GAATTTAATT TACCTAATCC TGATAATTGA CCCGCTTCAA ATAAGTNTGG TGCTACTACC 2280
GGTGGTTTAG GTAGTAAATA TTTAATCATC GCCTCAGTAT TTCCAGGATT AAGCACTTCA 2340
TCTTCGGCCT TTACAANCCA TGACAGGATC TTACTACAAA ATATTGTGCA ATTAGATGCT 2400
CATATACTTC TCCAGNACGG GTGGATATCT CCTCAAG 2437

```

## (2) INFORMATION FOR SEQ ID NO:192:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 479 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:192:

```

TTTCCCCNA CTCCATAATC TTGGCCAAAA TCATGTTTAA ATTTGATCTT GATTTGACC 60
CAAAGATAGC TTGCATATTG TTTCCCAACT TCTAANACAC CTGATGCGCC TAAAGCTTTA 120
AATACCTGCT ACATCTACTT TTGATTTATC CAACCCACTT CTACGCGCAG ACGTGTAATA 180
55 CATGCATCTA AATGTTTAAT GTTTTCTTTT CCACCCCATG GNATCTAAGA CATCANATGG 240
TAATTTTGCG AACTAGAGT TACGAATTTT AGTTTCTTCA TCTTCACGAC CTGGTGTGTT 300

```



# EP 0 841 394 A2

CAATTTAAAC TTACGAATTN CACAGTCGAA TGAGANGTAA NACACGATAG CAGACACGAC 360  
AGCGTTCTNN GACTGGTAAT NGGGCATNN GCCCAACCCA ATTTATGTGA GCTAACTCAG 420  
5 GTTGAACNTT GACCATCTTT NNATTCTTCA ACCCANAACC TGCAGGCGTA TTCATAANT 479

(2) INFORMATION FOR SEQ ID NO:193:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2593 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:193:

CCAAAGTGT CTGGCTAAG ATATTATCGC TAATCCGTGT TAAGCCACTC AGCTGTTTAA 60  
TATCAATTAC CTTTTNAAA TCCCTCCAAC CTAAATCTT GATTCCCACC TTGCTTTAAT 120  
25 AAAATCTTTC GTTTAGTTG TATTATGTTT AAACCTCTTG CCAGTCTTTT CGGTAAGATA 180  
ATCAATAATC TCTTATATG GGATGCGTGT CGGGTTTCCC GACAATATAT CTACTCTATT 240  
TATATTGTTA TTAATTGTAT TATTAATACT TGTATTGTCT CTTTAACATT TGTGATAATA 300  
GGGGTATTAA CAGAATTGTT AATAGGGGTA TTATCATTG TGTAAATAGG TCTTATCATT 360  
30 TCTGTTAAGG GGTATAGCTT TCTTTGTTTA ATTTCAATAC CATTTCTAAT GATTCAACA 420  
TGTAATATC CACATCTTT TAAGTTGGCT ATACGGCGTG ATACAGTAAC TTTTGTAAC 480  
TCATATAGTT TCGCAAAGTA ACCATTACTT GCTGTGCAGT ATCCGTACTT GTTACTTAAA 540  
GACGTTATTT CTGCAAAAAG TAACTTTTCG CTGTCAGTAA GTCGGTTATC GTATCTGACA 600  
35 TTTGCCGTTA TTATTGAGTA GTAACCTGGT TGTTCAGTCA TTCTCAGCAC CTCTTTTCAG 660  
TGCTTTGAAC TTGTCTGGTA TCTCCCAGTT AGATATGAAT TCTTTCAATT CATCAGTCAT 720  
AGGTACTTCG TTTAGTATTG CGTCATCACC AAACAGATAT AAAACTATCT TGTATATATG 780  
TAATGCCGCT TTTTCTGTGC TATCAAAATA TTCCGTAGTA TATAGCTTGC CATCAATTTT 840  
40 TTTCTAACA GAAATCGAT GCGGCATATC AGAGTGGTTT CCAATTCCCTC TGTAGCCATA 900  
TTTACTAGCG TTTCTGTGT TGTAAATATT TTTATGAGGA AAATAATCTC TAACTTTACG 960  
CATATCTCTT CCAATTATGT TTTTAAACC TTCCACCACC CCAAAACCTCA TCTACAGCTT 1020  
NGTTATACGC TTCAGCTGCT AATTCTTCTT GTTTNAAAAA TACCTAAACG TTTTNGGATT 1080  
45 TCCTTCAACA TTTATATTGG CTGTCCATTT ATTGCAATCT TNACGCCAGT TAACACCTTT 1140  
ANACTTAGAA GAACCATAC TTTTAGGNT CTCCCATCTT GANCNGTTNN NACCTTTTGT 1200  
AGTTAAATTA GATTTAGTAA AATTATTATT TTTAATTTT TGAAAACCTC CTCTTTTAT 1260  
AAAGGTTGTT AATAGTACAA CCTTCTCTC GTTGCCAATC CTAGTATGGA TATATCTTGT 1320  
50 ATTACCTTTA TAGTATTTGA ACCATTTATA TCTGTTCCAC TCTTTCATAA TCTTCATCAT 1380  
CAACAAAAAT TTCTTCTCCA TCTTGTAATA ATATCGATTT AACCATTATT CTCTTCCTTT 1440  
CAGCATTTTG TTGAGCCTCT CATCAACTTT TATCCACGAG TCATGCAAGT GATATTTATC 1500  
ATCAAAACGAC TTAACACCAA TCGCATGTTG CTCGTTGTGA TGTTGCGGAC ATAACGCTAA 1560  
55 TACATGTTTG TTGTAGTGAT TCATTTTGTT TCTGTTTATT CCTCTGCCGA CTGCTTCATA 1620  
ATGCGCTAGG TCTGCGTGAG GCTTCCACA AATTACACAG TTGCGGTTGA TTGTAGCCCA 1680

EP 0 841 394 A2

5 ATACAATAGT GCTTTATCCT CACTTAACAA CTTGCTCTCG TGCCTATGCT CATAGGTATT 1740  
 TGATGATGAA ACATAAACGC TATAATCAGT TCTATTAACCT CCCTTGCAAC TTTCATAGAA 1800  
 CAGTCGCGCA GACTGATTTC TTCATAACCT TTCATAATTT CCAATTCTGT TTGTAATAAT 1860  
 TTNCTAATTG ATTCCACCGG TTCTCCCCAG TGAAGTTCTA TATCTCTACA CATNGCGAAT 1920  
 ATTTTTTTGC GTTGTCTTAT AGATAGTTTT TTATTATCCG GAACCTCTAC TTCTGCTTTT 1980  
 AGTGGATATC CGTTTTCTAG TAAGTCAATG TGACTTTGTT CAAGTTCAAC ACCANTAGCA 2040  
 10 ACGACGGAAT AAGTNCCGTC ATTGTCNNTC TGGNATCTTG TAATGTATTG CANNTAAACC 2100  
 CACACCTTAA ACGCTAAATC TTGGTCGTCA TATCCAAATT GCCCCTGCT TTCAAATGGA 2160  
 TTGCTTTGTT GAGACATTGA TGTTTGTTGT TGTGCCCCGT TATTTTCTTC AGCTTTTTCG 2220  
 TTATCTGTCT TCGGAATAGG TTTGTTAACA ACATCATCGC CCTTTTGTGA AGGTTTAATA 2280  
 15 AATGAAAAAT CCGTAAAATA CTTACCTTCA TCTTCATTGA ATTTCCATTT CAATACCAAG 2340  
 TGACANAAC TACCAATAAG ATCATTGGTA TCANAATCTA AGCTAGGAAG ATTTAACTTA 2400  
 ATACCTAATC GAGTAACTAA TTCAATCAAT TGTTTTTCTT GGAAATCATA TTTATACGGC 2460  
 GGTACAAATT GATTATGTTT ATATTGTTTG CCTTCATCAT TTTCAAATAC GATTGTGAAA 2520  
 20 TATCTATTTT CTCTATCATT AGAAGAAAAC CTGATCATCT TTTTATAGGT TCTTAGATTG 2580  
 GAATTGTTGA GGG 2593

(2) INFORMATION FOR SEQ ID NO:194:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3191 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:194:

40 CCCCAGTCT CATATGAAGT AGTTGGATTA GTTCCGAAGT ATGAGGATGG GAATTTCTTG 60  
 ACATAGTAAA GTGCCGGATG ATAGTGCAAG AGGTATCGTT GCACAAAATA TGACAACTGC 120  
 TGAGATTGAA ACATTTGGTT CAGATCCAGT TATTATGCCA ACGGGGTGCC CCTGGTATTA 180  
 TTTTCGGTAA AACAACAAAC TCAATGATTA ATACAGGATC AGCGGCTTCC ATTGTTTACC 240  
 AACAAAGGCGC TATTTATGCT AATGCTGAGT TCATTCAAAT TCATCCTACT GCAATCCCTG 300  
 45 GTGATGATAA ACTCCGACTA ATGAGTGATC AGCACCNTTG TTGAAGGTGG ACGAATTTGG 360  
 ACATATAAAG ATGGTAAGCC TTGGTACTTC TNAAGAAGAG AAATATCCTG ATTATGGTAA 420  
 CTAGTACCT CGTGATATCG CAACCGCGTG ANATTTTCGA TGTATGTATT AACCAACAAA 480  
 TTAGGNATAA ATGNCGAAAA CATGGNATAT CTTGATTGTT CACATAAAGA TCCAACATGA 540  
 50 GTTAGATGTA AAAC TAGGTG GTATCATTTGA AATTTATGAN AAATTCACCTG GTGATGACCC 600  
 ACGCAAAGTA CCAATGAAGA TTTTCCCAGC TGTTCACCTAT TCAATGGGTG GTCTATATGT 660  
 AGATTATGAT CAAATGACAA ATATTAAAGG GTTATTTGCA GCTGGAGAAT GTGACTTCTC 720  
 TCAACATGGT GGTAACCGCT TAGGTGCCAA TTCATTGTTA TCAGCGATTT ATGGTGGTAC 780  
 55 AGTAGCAGGT CCAAACGCGA TTGATTAATA TTTCAAATAT TGATCCGATC CATATACTGA 840  
 TATGGACNGA AGTATTTTTG AAAAACGTAA AGCTGAAGAG CAAGGAACGT TTTGATAAAT 900

EP 0 841 394 A2

	TATTAGCTAT	GCGCGGTACA	AGAAAATGCC	TATAAATTAC	ACCGTGAAC	TGGTGAAAT	960
	ATGACACCAA	ATGTAACGT	TGTTTCGTGAA	AATGAAAAAC	TGTTAGAAAC	AGATNAAAAG	1020
5	ATTGTTGAAT	TGATGAAACG	TTATGAAGAT	ATTGATATGG	AAGATACTCA	AACTTGGAGT	1080
	AACCAAGCGG	TATTCCTTAC	CCGTCAACTA	TGGAACATGT	TAGTACTTGC	ACGTGTTATT	1140
	ACGATTGGTG	CATATAACCG	TAACGAATCA	CGCGGTGCCC	ATTATAAACC	AGAATTCCCTG	1200
	ATAAGTATTT	TCGTCAGAAG	TACATAATAT	TTAAATCATT	TAATTTTGAG	AAGCATGGAA	1260
10	AATTTTGGA	CAAATGGTTT	TACGTAAGAA	AATGGAAACA	TAAGATTTTA	GATGGTCATC	1320
	AGCTTAATCA	AAATATATAT	GATCAGCGTC	ATTTAATGAC	AATCAATACT	GATGAAATTG	1380
	AAAAAATGAT	TATAGAGACA	AAGAGGGCAG	AGTTGATTCA	TTGGATATCG	ATACTTCCAG	1440
	TCATCATATT	CAATAAAGGC	TCTCGTTTAG	TAAAGTATAT	AAATATTTTC	TATGCAATGA	1500
15	TAGCTAATGT	TCCAATCATT	ATTGTGCAAC	GCTATAATCG	ACGAGATTAA	CGCAGGTACT	1560
	ACGCATTTTA	AAACGAAGAG	GTGAACGTCA	TGACTAAACA	TATCATCCGN	TATTGGGTGG	1620
	GNGGCTTAGG	TGGGATTTCN	TGCAGCAATT	CGAATGGACA	AAAGTGGCTA	TTCCGNCTCA	1680
	TTATATGAAC	AAAATACTCA	TATAGGAGGC	AAAGTGAATC	GNCATGAATC	AGATGGCTTT	1740
20	GGCTTTGATT	TAGGTCCATC	TATTTTAACG	ATGCCTTATA	TTTGTGAAAA	ATTATTCGAA	1800
	TATAGCAAGA	AGCAAAATGTC	AGACTACGTT	ACAATCAAGC	GTTTGNCACA	TCAATGGCGT	1860
	AGCTTTTTTC	CAGATGGCAC	GACTATCGAT	TTGTATGAAG	GTATTAAAGA	AACAGGTCAG	1920
	CATAATGCGA	TATTGTGCGA	ACAGGATATA	GAGGAACTGC	AAAATTATTT	GAATTATACA	1980
25	AGACGAATCG	ATCGTATTAC	TGAAAAAGGG	TATTTTAACT	ATGGTTTAGA	TACACTATCT	2040
	CAAATTATTA	AATTCATGG	GCCATTAAAT	GCTCTTATTA	ATTATGATTA	TGTACATACT	2100
	ATGCAACAGG	CCATAGACAA	GCGTATCTCG	AATCCATACT	TGCGACAAAT	GTTAGGCTAT	2160
	TTTATCAAAT	ATGTAGGTTT	TTCATCATAC	GATGCGNCAG	CTGTATTATC	TATGTTATTTC	2220
30	CATATGCAAC	AAGAGCAAGG	CCNTTGNTAT	GTAGAAGGTG	GAATCCATCA	TTTNGCCAAT	2280
	GCCTTGGA	AGCTAGCGCG	TGAAGAAGGT	GTCACAATTC	ATACAGGTGC	ACGTGTGGAC	2340
	AATATTAAAA	CATATCAAAG	ACGTGTGACG	GGTGTGAGAT	TAGATACAGG	TGAGTTTGTA	2400
	AAGGCAGATT	ATATTATTTT	AAATATGGAA	GTCATACCTA	CTTATAAATA	TTTAATTCAC	2460
35	CTTGGATACT	CAACGATTAA	ACAAATTAGA	GAGGGAATTT	GAGCCGGCAA	GCTCAGGATA	2520
	TGTGATGCAT	TTAGGTGTTG	CTTGCCAATA	CCCGCAATTA	GCACATCATA	ATTTCTTTTT	2580
	TACGGAAAA	GCTTATCTCA	ATTATCAACA	AGTTTTCAT	GAAAAGGTAT	TGCCAGATGA	2640
	TCCGACCATT	TATCTAGTAA	ATACGAATAA	AACATGATCAC	ACACAAGCGC	CAGTAGGGTT	2700
40	ATGAAAAATAT	CAAAGTCTTA	CCACATATTC	CATATATTCA	AGATCAGCCT	TTTACCACTG	2760
	AAGATTATGC	GAAGTTTAGG	GATAAAATTT	TGGATAAATT	AGAAAAAATG	GGACTTACTG	2820
	ATTTAAGAAA	ACACATTATT	TATGAAGATG	TTTGGACACC	GGAGGATATT	GAAAAAAATT	2880
	ATCGNTCTAA	TCGTGGTGCA	ATATATGGTG	TTGTGCGAGA	TAAAAAGAAA	AACAAAGGAT	2940
45	TTACCTTTTC	TAAAGAAAGT	CAGTATTTTG	AAAACCTGTA	CTTTGTAGGT	GGATCAGNAA	3000
	ATCCTGGTGG	TGGCATGCCA	ATGGTTACAT	TAAGTGGGCA	ACAANTCGCA	GACAAANTNA	3060
	ACGCGCGAAG	AAGCGAANGA	ATAGGANGGT	GANATCTATT	GAAATCGGTN	NTCACGACTA	3120
	TTTAANAGAC	AATAGTGGNN	ACCCATCNTC	TTNTGGGGNC	TTGGGAGCAT	TTGGATNTTC	3180
50	NNCGGGGGNC	A					3191

(2) INFORMATION FOR SEQ ID NO:195:

55 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1189 base pairs

EP 0 841 394 A2

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:195:

CCTGGAATAA	TTGGTGGTCC	GTAATTGATG	GGCCAAGTTG	AAATTGGAAG	CTAAAAAAGT	60
CTTAATCCGC	CCAAATTGTT	AACAACACAC	CTATTGACTT	TGTCAATCCA	ACCCATCAAT	120
CCTTCATTAT	TAACATGTCC	CAGGTATCTT	TAAGNTTGAT	GAAACATTAG	ATACCAGTTG	180
TTGGCGAAGG	CTTAAAGCTT	TGCCANCTTG	AAAAATTAAG	CCGGAATTAA	GATGCATTGC	240
GTCGCCTATT	TGAATCCAAA	CCGATAGTGT	TAAATATGAT	AAATTAAAAG	CACGTTATGA	300
GCGTTTCCAA	AATCAATCAT	TCAAAAACTT	AGATTATGAT	TTCGAAAGCG	TGCGTACTTC	360
AAGACAATCA	CCATTGCGCG	AACGTATTGA	ACAACAACAA	AAACGTTTGA	ACTTACCGGA	420
TTTACCAACA	ACAACTATTG	GATCATTCCC	ACAAAGCCGA	GAAGTTCGAA	AATACCGTGC	480
AGATTGGAAG	AACAACGCA	TTACAGACGA	AGCATATGAA	ACATTCTTAA	AAAATGAAAT	540
TGCTCGATGG	ATTTAAATTC	AAGAAGACAT	CGGCTTAGAT	GTATTAGTTC	ACGGTGAATT	600
TGAACGTAAT	GACATGGTTG	AATTCCTCGG	AGAAAAATTA	CAAGGTTTCT	TAGTAACTAA	660
ATTCGGTTGG	GTGCAATCAT	ATGGTTCACG	CGCCGTAAAA	CCACCAATCA	TTTATGGTGA	720
TGTAAATGG	ACAGCGCCTT	TAACGTGTGA	TGAAACAGTT	TATGCACAAA	GCTTAACAGA	780
TAAACCAGTT	AAAGGTATGT	TAACGTGACC	TGTAACAATT	CTAAACTGGT	CATTGAACG	840
TGTTGATTTA	CCACGTAAAG	TCGCTCAAGA	TCAAATTGCT	TTAGCAATCA	ACGAAGAAGT	900
ATTAGCACTT	GAAGCTGCAG	GAATCAAAGT	TATCCAAGTT	GACGAACCTG	CATTACGTGA	960
AGGCTTACCA	TTACGCTCTG	AATATCACGA	ACAATATCTT	AAAGATGCTG	GTTTTATCAT	1020
TTAANCTTGC	AACGTCTTCA	AGTTCGGTGA	TGAANCTCAA	ATCCATACAC	ATATTGTGTT	1080
ATTCTCAANT	CGGGCAAATC	AATCAATGCT	ATTCAAGATT	TAGATTGCTT	GATGTTATTC	1140
AATTGAACAC	TCCGNGCCNT	GGGGATTTAA	TTAAGNTTGG	AGGTTTAT		1189

(2) INFORMATION FOR SEQ ID NO:196:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3815 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:196:

GGAGCATAAT	AAAGAGCGTA	TCTTAAACAT	TAATTTTAAG	TGTGATCCCA	AAGTAATTTT	60
CGTGATTCAT	TGAAAATTGT	TCCGAAAGNA	CCGTTGTTCA	TCCAAGACAA	TATTTACCTT	120
AACAATATPG	ATTGGGGGTT	ACCTGGAACG	ATGTTATTAA	GTTTGGTTAG	CTGACAAAAA	180

EP 0 841 394 A2

	TAGTCGTTAC	CGCCATCGCA	GGACCCAATG	CCGATTGTAA	TGGTTAGCAT	TATTAAATCA	240
	ATGATGCATC	CAAAAAATGG	AATATATAGA	TGCAAGTCAT	ATTGTGATTG	CGTTGTTAAT	300
5	CATCCTTGTG	ATAATCACAT	TGTTTATCTT	ATTATTCAAT	GAATTAGTAG	AAGTGAGAAT	360
	ACCATATATC	GATTTAATGA	ACGTTTCCGC	AACAAATATG	AAATCTTATT	TATCTTGGAA	420
	AGTTAACCCT	GCAGGCAGTA	TTACTTTAAT	GATGAGTATT	TCAGCATTG	TTTTCTTGAA	480
	AAGTGGCATT	CATTTTATTT	TATCTATGTT	TAATAAAAGC	ATATCAGATG	ACATGCCAAT	540
10	GCTGACATTT	GATAGTCCAG	TAGGTATTTT	AGTATATTTA	GTGATCAAAA	TGTTATTGGG	600
	TTATTTTTTA	TCGAGATTTT	TAATCAATAC	TAAACAAAAA	TCCAAAGATT	TCTTAAAGAG	660
	TGGCAATTAT	TTTTTCAGGAG	TTAAACCTGG	TAAGGATACA	GAACGTTATT	TAAATTATCA	720
	AGCAAGACGC	GTATGTTGGT	TTGGATCGGC	ATTAGTTACA	GTCATTATTG	GTATACCGCT	780
15	TTATTTTACA	TTGTTTGTAC	CGCATTTATC	TACTGAAATT	TATTTCTCAG	TACAACTGAT	840
	TGTATTAGTT	TACATCAGTA	TTAATATTGC	AGAAACAATT	CGTACATATT	TATATTTTGA	900
	TAAATATAAG	CCATTTTTTA	ACCAGTATTG	GTAAGGAGGT	AATTATGAAA	TACTTTATTC	960
	CAGCTTG GTA	CGATGACCAA	CGATGGTGGC	AAGACACGAC	TGTGCCGTAT	TATCAACTAC	1020
20	AAAATAAGAC	GGAATTTGAC	GATATGATTA	GTTTAATGGG	AATGCACCTT	GAAAAAGACT	1080
	TAGATTATCA	ACTGATTGTT	CTCAATCATG	CACCAAATTT	AAGAACATTT	TTACATCGAT	1140
	ATGACTTATA	TGAAACAAAG	TATTCGTCTG	TGTTTGATGA	AATTCAAGGA	TTCAGTCACC	1200
	ATGCGCCACA	AGCGATTAAT	TATCATCACT	TAAATGGNC	GGATGATGTT	GGAGTNTGGG	1260
25	TACACGCCCG	NATTTATTAA	AAATGTGTGA	CGAGTGAACA	GACCTATACA	NATATTTATT	1320
	TNAGNCAAGA	AGGGTATTCA	NTTGGGTTTG	NAGANTGATT	NGAAAGAGAT	CAAGTNACAA	1380
	CGGCGTTATA	TTTTTGATGA	CAGAGGTTAT	TTATCAGCGA	TACGTTATTT	TGATGATCAG	1440
	GGAGAGGCTT	CTTACCAACA	ATATTTAACG	GATTAATGGA	GATTGTGTAA	CTTCATGGAA	1500
30	GATTGGANAA	ATGGCAGGAG	TCACTGTATC	NANAAGGATA	TTCAACATCA	CTATCAACAA	1560
	ACAGAATATA	ACAATATGGC	TCAACTAATT	GGAAGAAAAA	TTTCAAGCAA	TGATTGCACA	1620
	ACAAATACAT	GAAGATGATC	ATGTGATTGT	GGCTTCAGAT	GCTAGGCACA	ATCGACAAAT	1680
	AGCCAATCAT	ATTCCAGCGA	AATTGTAAAG	TTATTCATTT	TTTAAAAATA	GAAATGAANC	1740
35	TGTGTCAGAT	GAGGAATATC	AATCTATCGT	AAAGAATGCC	CATTTAATTG	TTGATAGTGT	1800
	GCAACTAGAA	CGTGATTTAA	TTAGTCATCA	AGAGAAGTAT	CAGCGGGAGA	ATACAATGAT	1860
	TCGAATCACA	CCATTTGAAA	CGAGACAATC	ACCTAATATA	AGTAGTCAAT	TGATGGAAAC	1920
	ATTTATAGGT	GTATGGATAG	ATGGTATGAG	TGACGCTGAT	TTGCAACAAA	TGATGCAACG	1980
40	ACTTGTGGAT	TATATCGCAC	AGGAAGATTA	TTACCGTTTA	ATTTTATTAT	CGCGCCATCA	2040
	AAATGACATA	CCGATGTGGC	TTCTGTGAATG	TATTACGTCG	GTAAATGAGG	AATACCAAGC	2100
	TAAACAGAAT	GCGGATGTTA	ATGTTTCAGC	ATTAATGACA	CCTGAAGATC	AAGATGACAT	2160
	CATTGCTGTT	AAGACGATAC	ATGCTGAACA	TGATGTTGTA	GAAGCATTGC	GGACGTTGCG	2220
45	ACTTGTGATA	GATATGTCAA	AAGAACCTGA	TTTGATTTTA	CAAATTAGTG	CAATTAGCGC	2280
	TGGGATTCCA	CAAATTAATG	GTCAACAAAC	AGATTACGTC	TCTGATTATG	ACAATGGCCG	2340
	TATTATAAAT	ACAGCTGGAT	GAATTAGATG	ATGCGTTAAA	TTATTATTTA	TTTTATTTGA	2400
	AAAATTGGAA	TTATGCGTAC	GCCTATTCTT	TAAAATTAAT	AGATGCATAT	GCTTCTAAGA	2460
50	ATATTATTAA	TCAGCTCGAT	GAGTTAATAG	AAGGTGAAAA	TGATGCCACG	TAAATTTAGA	2520
	GTTTTGCAAA	TTGGAGGAGA	CGATTTAGAA	CCTATTTTTC	AACACAAAAA	AGGTGTGAGT	2580
	TGGGATTACT	TCGATATTGG	ATTGTTTGAA	TTTGATAGTG	GTTATGTAGA	GGCTATTGAA	2640
	GCGATTGTTG	AAGCAGAAGG	GCGCTTTGAT	TTTATCTATA	TTCAAGCACC	ATACTCGGAG	2700
55	ACATTAAACGA	ATTTATTACA	AATGATAAGC	GAACCATACA	ATACGTATGT	TGATGAATCA	2760
	TTTTGGTCAG	TTGAATATGA	ACAAGACGAA	AATTGTCCAA	AAATACGTTG	TTCAACCAAT	2820

EP 0 841 394 A2

5 TACATTACCG GAATATTGGA AGGAACGTAA TAATNAATTA GAGGCAGGNT AGCTTCTCAA 2880  
 GGACAAGGAT TGGAGATAAA GTTTCCTTA AGTTAGCACT TGTGCATCCG AATTTTAAAG 2940  
 GAGATGTCGA ATACCCAAGG TAATTCAAGA GCTCACGTTG AGTGGAGAAT TTGAAAAGA 3000  
 ATTTAAACCT ATCGCATCTT GNCAAAATA TCTCGTTTAC GATAAAGATA AAGTCATTCA 3060  
 NATATGGCCA GAATTTGATA TTGATGGTGC GGTGAGTTG CAATATACAT TTAGATTGAT 3120  
 TCAGACTGGC GCTGATGGTG CATTAAATGA ACAAATCATA TTAAGTATG ATATGTTAGA 3180  
 10 CAGTCCTTTA GAGATACCTG CGAAACCATT TGATGCTTAT ATAAGTGTA CTGTTAAGGC 3240  
 GCGTGGGAAC GGGACGGTAC ATTTAGGACC TAAACACACA CGATGGTCCA GANTAGANAN 3300  
 GGNTCAATTT TTACNTGGTG GGAATCGTTT CGAAGATANN CNCCGNCNGG AATTTAATTA 3360  
 TTAATCCNC CCTGGTGATA TGANACCCC ACTAAACGTA AATTTTAGTG GTTATCGANC 3420  
 15 ACCGGAAGGT TTCGAAGGAT ATTATATGAT GANACGTATG AATGCNCCGT TTTTACTTAT 3480  
 CGCCGATCCT CGTGTGANG GTGGTAGCTT TTATATCGGT TCATCTGAAT ATGAACAACG 3540  
 TGATTATCAA TGTATTGAC GAGACAATAT TCATGCTGTA AATCCAACAT CAGGTGCAGC 3600  
 TGGTAAGTAT CAATTCTTAC AATCAACTTG GGATTCAGTA GCACCTGCTA AATATNAAGG 3660  
 20 TGTATACCA GCAATGCTC CTGGAAGTGT TCAAGATGCC GCAGCAGTAA AATTATATAA 3720  
 CACTGGTGGC GCTGGACATT GGGTTACTGC ATAAGCCATT TATGCATAGC TAATCAATAG 3780  
 TTATATAAGT AACTTTTAGA TCGGAATATA TCGGG 3815

(2) INFORMATION FOR SEQ ID NO:197:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2848 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:197:

40 GGCNCGAGAC GCTTAATTAT GGAACNCCAA GATGCNATTG GACNCGGTTA TAACATTGGN 60  
 ACAGGCACCT TTAATAATTT ATTAGAGGTT TATCGTATTA TTGGTGAATT AAATATGGNA 120  
 AATCCAATCG AGCNTTGAAT TTAAAGGAAG CACGAAAAGG AGATATTAAC CATTCTTATT 180  
 GNAGATATCT CTAACCTAAA GGCATTAGGA TTTTGTGCTC CTAAAAAATA CAGTAGAAAC 240  
 45 AGGTTTAAAG GATTACTTTA ATTTTGAGGT AGATAATATT GAAGAAGTTA CAGCTAAAGA 300  
 AGTGGAATG TCGTGAATAT GACATTGAAG CTGTCCATAA TAATAAGGGT TATGCCATATC 360  
 AAAGAAAATT AGATGCACTC GAAGAAGTGA GAAAAGGCTA TTACCAATT AAACGTGCGA 420  
 TTGACTTAGT ATTAAGTATC GTTTTATTAT TTTTAACATT TCCGATTATG TTCATATTCTG 480  
 50 CCATTGCTAT CGTCATAGAT TCGCCAGGAA ACCCTATTTA TAGTCAGGTT AGAGTTGGGA 540  
 AGATGGGTAA ATTAATTAAA ATATACAAAT TACGTTTCGAT GTGTAAAAAC GCAGAGAAAA 600  
 ATGGTGCACA ATGGGCTGAT AAAGATGATG ATCGTATAAC AAATGTCGGG AAGTTTATTC 660  
 GTAAACACG CATTGATGAA TTACCACACC TAATTAATGT TGTTAAAGGG GAAATGAGTT 720  
 55 TTATTGGACC ACGCCCGGAA CGTCCGAAT TTGTAGAATT ATTTAGTTCA GAAGTGATAG 780  
 GTTTCGAGCA AAGATGTCTT GTTACACCAG GGTAAACAGG ACTTGCGCAA ATTCAAGGTG 840

EP 0 841 394 A2

```

GATATGACTT AACACCGCAA CACAACTGA AATATGACAT GAAATATATA CATAAAGGTA 900
GTTTAATGAT GGAACATATAT ATATCAATTA GAACATTGAT GGTGTTATT ACAGGGGAAG 960
5 GCTCAAGGTA GTCTTAATTT ACTTAATAAG TTCAAATAAA AGTTATATTT TAAAGATTGT 1020
GACCAATNGT TACAGTNTAA CGAGGANTCC CTTGNGACNG TATCAAATGG CATTAAAGAA 1080
ATATGTNCCA TCNTTTGATT TGCNTGGCCA ATAAATACTA TTCATCTTGA TGAGATAAGC 1140
CATGTTAAGA AATTGAAAGT ATAGCATTA NGGGGTTTGT AACAGTTGAA AATTATATAT 1200
10 TGTATTACTA AAGCAGACAA TGGTGGTGCA CAAAACACAT CTCATTCAAC TCGCCAACCA 1260
TTTTTGCCTA CACCATGATG TTTATGTCAT TGTAGGCAAT CATGGACCAA TGATTGAACA 1320
NCTAGATGCA AGAGTTAATG TAATTATTCT CGAACATTTA GTAGGTCCAA TTGACTTTAA 1380
ACAAGATATT TTAGCTGTCA AAGTGTTAGC ACAGTTATTC TCGAAAATTA AGCCTGATGT 1440
15 TATCCATTTA CATTCTTCCA AAGCTGGAAC GGTGCGACGA ATTGCGAAGT TCATTTTCGAA 1500
ATCGAAAGAC ACACGTGTAG TTTTACTGCG GCATGGATGG GCTTTTACAG AGGGTGTAA 1560
ACCAGCTAAA AAATTTCTAT ATCTAGTTAT CGAAAAATTA ATGTCATGTA TTACAGATAG 1620
CATTATTTGT GTTTCAGATT TCGATAAACA GTTAGCGTTA AAATATCGAT TTAATCGATT 1680
20 GAAATTAACC ACAATACATA ATGGTATTGC AGATGTTCCC GCTGTTAANC AAACGCTAAA 1740
AAGCCAATCA CATAACAATA TTGGCGAAGT AGTTGGANTG TTGCCTAATA AACAAGATT 1800
ACAGATTAAT GCCCGACAA AGCATCAATT TGTATGATT GCAAGATTTG CTTATCCAAA 1860
ATTGCCNCAA AATCTAATCG CGGCAATAGA GATATTGAAA TTACATAACA GTAATCATGC 1920
25 GCATTTTACA TTTATAGGCG ATGGACCTAC ATTAAATGAT TGTCAGCCAC CAAGTTGTAC 1980
AAGCTGGGTT AAGAAAATGA TGTCACATTT TTGGGCAATG TCATTAATGC GAGTCATTTA 2040
TTATCACAAAT ACGATACGTT TATTTTAATA AGTAAGCATG AAGGTTTGCC AATTAGCATT 2100
ATAGAAGCTA TGCTACAGG TTTGCCGTGT ATAGCCAGTC ATGTTGGCGG TATTTTCAGAA 2160
30 TTAGTAGCTG ATATGGTAT ATGTATGATG AACACCCAAC CCGAAACTAT TGCTAAAGTC 2220
CTGGAAAAAT ATTTAATAGA CAGTGATTAC ATCAAAATGA GTAATCAATC TAGAAAACGT 2280
TATTTAGAAT GTTTTACTGA GGAGAAAATG ATTAAGAAG TGGAAGACGT TTATAATGGA 2340
AAATCAACAC AATAGTAAAT TACTAACATT GTTACTTATC GGTTTAGCGG TTTTATTCA 2400
35 GCAATCTTCG GTTATTGCCG GTGTGAATGT TCTATAGCT GACTTTATCA CATTACTAAT 2460
ATTAGTTTAT TTACTGTTTT TCGCTAACCA TTTATTAAAG GCAAATCATT TTTTACAGTT 2520
TTTCATTATT TTGTATACAT ATCGTATGAT TATTACGCTT TGTTTGCTAT TTTTGTATGA 2580
TTTGATATTT ATTACGGTTA AGGAAGTTCT TGCATCTACA GNTAAATATG CATTTGTAGT 2640
40 CATTTATTTT TATTAGGGA TGATCATCTT TAAGTTAGGT AATAGCAAAA AAGTGATCGT 2700
TACCTCTTAT ATTATAAGCA GTGTGACTAT AGGTCTATTT TGTATTATAG CTGGTTTGAA 2760
CAAGTCCCTT TACTAATGA AATCGTTATA TTTTGATGAA ATACGTTCAA AAGGATTAAG 2820
GAATGACCCT AACTATTTTC CTCGTGCC 2848

```

(2) INFORMATION FOR SEQ ID NO:198:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2144 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:198:

5	CCCCAAAGAA ATACANTTGT TGGTAGAAGA ACCCCCAACC CACAAGNTGG GGACCCTTCA	60
	TTTCCAACCN TATGGCGGAG GGTAAAGTA AATTCTTGAA GCTTTGGTTG CCCNCAAATA	120
	ATATGATATT TCCGAAATGG GCCGGAGCCG TATCCAACAA TTAAATCT GCCGGAGGTA	180
10	TTAAAGCATT ACTTAACCGC TCCGATTAAG ATGNNATTGA AGGAGAGATT AATATGGTAG	240
	AATCAATGCT AACTTTTATG CTTGGGCCAT TAAGACAAAT CNCTGATTTT TATATGGAAC	300
	ATTTACTCGT AAGTAATTCC ATCGTCATTG CAGGTTATTT TGCGACAGGT ATTTTAAAA	360
	AGAAAAAGT TGTGAATTAA ATCANATTTG AGGTGATTTA CAAGTGAAAG CATTGAAATT	420
15	ATATGGCGTA GAAGATTTAC GGTATGAGGA TAATGAAAAG CCAGTCATTG AAAGTGCAG	480
	TGACGTTATT GTTAAAGTAC GAGCGACTGG CATATGTGGT TCAGACACGT CACGAAACAN	540
	AAANATGGGG CCATACATTA AAGGTATGCC ATTTGGTCAT GAATTTTCAG GTGTAGTAGA	600
	TGCCATTGGA AGTGATGTTA CGCATGTTAA TGTGGCGCAC AAAGTGACAG GTTGCCACAG	660
20	AATACCTTGT TATCAATGCG AGTAGTGT TTGAAAGGTGAA TATGCACGAT GTGAAAAAGT	720
	TATTCGTCAT TTGNCTCAAT ATGAACCTTG GATCGTTCNG CGGAATATGT CAAATCCCA	780
	CCCCAAAATG TTTTNAAGGG TTCCAAGCCA ATGTTGATTA CCATGAANC AACCAATGGT	840
	TTGAGCCATC AGCCCGTTGT TGCGCATGGG TTTTATAAAT CCGAATANAC AACCTGGTAT	900
25	GACTGTTGCA GTAATGGGGT GTGGCAGTAT AGGTTTGTTA GCTATTCAAT GGGCACGAAT	960
	ATTTGGTGCT GCACATATCA TCGCTATAGA TATAGATGCG CATAAACTAG ATATTGCAAC	1020
	ATCATTGGGC GCACATCAAA CAATCAATTC AAAAGAAGAA AATCTTGAGA AATTCATCGA	1080
	AAATCATTAC GCCAATCAAA TCGATTAGC TATAGAATCA TCAGGTGCTA AAGTTACGAT	1140
30	TGGTCAAATA TTGACGCTAC CTAAGAAAGG TGGCGAGGTG GTATTACTCG GAATACCATA	1200
	TGATGATATT GAGATTGATC GCGTTCATTT TGAAAAAATT CTGCGTAACG AGTTGACAGT	1260
	ATGTGGCTCT TGGAAGTGT TGTCCAGTAA TTTTCCGGGC AAAGAGTGGA CGGCAACCTT	1320
	ACATTATATG AAGACGAAAG ATATTAATGT AAAGCCTATT ATTTCTCAT TTTTACCGTT	1380
35	AAGAAAAAGG CCCGGAGACA TTTTGATAAA TTAGTTAATA AGAAAGACCG ATTTGATAAA	1440
	GTCATGTTTA CGATTTATTA GTATGCACCT TTGAGGACGA AAACGCTGGT ATAGTTATAG	1500
	CTATGAAAGT GCGAATGCCG TCTGGTCTAC AGATACTATC GAAATAATTC ATCTTCGAAT	1560
	ATACGTTGAT AAATAGCCCG TTTACTTGTG TGAAATATGC TTGTGAATCG GTTGTTTTC	1620
40	ATTTTGATATA CTAAATGA GATGGCAATA TTTGATAATT TTTAAAGTGA AAATCAAGTA	1680
	CAGCCACTTA ATAAGATAAA TTTATTATAA TATATGGTAA AATGATGGCA GTAATAATGA	1740
	ATTTGAAAAA GAGTAAACAT TAATACCTTT AACAATTTAA TATCGTCAGA GTTAATGATT	1800
	AACTGCATGG CAAAACAAC TAGAATGGTC AGTTACNNAA ATACATTTT ATAAAAAATT	1860
45	ATCACACTAT TGTGACAAC ATCTTTGGAT TAATAAAGA GGCAAGTGAG CAATAGGTTA	1920
	GGCTTATGTG CGGGCATAGG TCAGTAATGT AGAAATGGAA ATGATGTAAT ACGTTAATTC	1980
	GTTTGAATCT TTAAGATAT TGGATAAATC AGATAATCGC CTGTTATACG CATCAATCAA	2040
	TGTTTTATAC ATAAATATGT CTGATATTT TAAGTCGTTA TTTTGAATG GAATAAATGG	2100
50	AATCTTACCC CACGACCTG TACTAAAATG CGTTTTTTRD CTNS	2144

(2) INFORMATION FOR SEQ ID NO:199:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2178 base pairs



EP 0 841 394 A2

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:199:

CAAATAACGG	ACCCTTTNGT	TAATATCCCA	TNTTTTTGGC	GGAGTATGGG	ACCTTAGCAG	60
AGAAAATAGC	CATAAGATAA	GATTTTCGCC	CANNCGTTGG	AGCCCAACTT	ACATTGTTTG	120
TAGAATTTCT	TTTCGAAATT	CTTTGTGTTG	GGCCCCCACC	CCAACTTGCA	TCGTCTGNAG	180
AAATTGGGGG	TCCAATTTCA	CTGTGTTGGG	NCCCCCTGAA	CTAAATAGTA	TGGAAATATA	240
TATATCTAAA	TAATTAAAAC	GGNCAATAG	TTTAATATGG	CAGTAACAAT	TTTTATAATT	300
TAGATATTTT	GATATAGACT	AATGCTTAAA	TTAAGCAACT	TTTCTANNNA	GGGAGTATGA	360
AAATGATTAG	TACTTTGAAT	GAAATATGA	AATGTATCGG	AAGACCNCGA	TACAATTATT	420
AATACACAAG	ACATGTTAGA	CCAGATCCAG	ATGCATATGG	TTCAACAACCT	GGTTTAAAAAT	480
ACTATATTTCA	GCAGAAATTT	CCGCAAAAGC	AAGTATTTGC	AGTGGGTGAA	GCGGAATCAT	540
CATTAAGTTT	TATGGAGAG	TTGGATAACA	TTGATGATAA	AACATATCAA	GATGCGCTTG	600
TAATTGTATG	TGATACTGCC	AATGCTCCAC	GAATTGACGA	TGAACGATAT	AGTACAGGTA	660
GTAAACTTAT	TAAAATAGAT	CATCATCCTG	CAGTTGATCA	GTATGGTGAT	ATTAATTTAG	720
TTAATACGAA	CGCGTCATCT	ACAAGTGAAA	TCATTTATGA	TTTAATCTCA	CATTTTAATG	780
ATGAAGCAAT	TGTTAATAAA	GGCACGAGCG	AGTGTTTTAT	ACCTTGGTAT	CGTCGGTGAT	840
ACTGGGCGAT	TCCTTTTTTAA	CAATACCTCA	GAACATACTA	TGGAAATTGC	TGGAAAGTTA	900
ATTGGGCATG	ATATTGATCA	TAATGCGCTA	TTAAATAAAA	TGATGGAGAA	GGACCCAAAA	960
ATGTTGCCGT	TTCAAGGTTA	TGTTTTACAA	CATTTTCGAAC	TTATGGATGA	TGGATTCTGC	1020
CAAGTTAAAA	TAAGTGAAGA	TGTATTGGAG	CAATTCGGTA	TTCAAGCCAAA	TGAAGCATCT	1080
CAGTTTGTTA	ATACAATTGC	TGACATCAAA	GGTTTGAAAA	TATGGGTGTT	TCCAAGTTGT	1140
TGAGGGTACT	GAAATAAGAT	GTGCTTACG	CTCTAAAGGG	CAATTGATTA	TTAATGATAT	1200
TGCGCAAGAT	TTTGGTGGCG	GTGCCCATCC	GGAATGCGTC	AGGAGTTTCA	GTGAACAGCT	1260
GGGATGAATT	TGAGCAACTT	GCTACAGCTT	TACGCACAAA	ACTTAACTAA	TAGAAAGGAG	1320
CCATTCAATC	ATGGTGGCAT	ATTTAAATAT	TCATACGGCT	TATGATTTGT	TAAATTCAAG	1380
CTTAAAAATA	GAAGATGCCG	TAAGACTTGC	TGTGTCTGAA	AATGTTGATG	CACTTGCCAT	1440
AACTGACACC	AATGTATTGT	ATGGTTTTC	TAAATTTTAT	GATGCATGTA	TAGCAAATAA	1500
CATTAAACCG	ATTTTGGTA	TGACAATATA	TGTGACAAAT	GGATTAAATA	CAGTCGAAAC	1560
AGTTGTTCTA	GCTAAAAATA	ATGATGGATT	AAAAGATTG	TATCAACTAT	CATCGGAAAT	1620
AAAAATGAAA	ACAAATGAAA	ATGTTTCATT	TGAACTATTG	CAACAATTTT	CATCGAATTT	1680
GATTATCATT	TTTAAAAATG	TTGCAGACGA	ACATCGTGAC	ATTGTTCAAG	TTTGTGATTC	1740
GCATGAAGAT	ACGTATTTAG	ATCATCAAAG	TGTTTTAGTT	CAGGGTATAA	AGCACGTATG	1800
GATTCAAAAT	GTTTGTTACC	AAACACGTCA	AGATGCCGAT	ACGATTCTTG	CATTAGCAGC	1860
TATTAGAGAC	AATGCAAAAT	TAGACTTAAT	TCATGATCAA	GAAGATTTTG	GTGCACATTT	1920
TTTAAC TGAA	AAGGAAATTA	AACAATTAGA	TATTAACCAA	GAATATTTAA	CGCAGGTTGA	1980
TGTTATAGCT	CAAAAGTGTA	ATGCAGAATT	AAAATATCAT	CAATCTCTAC	TTCCTCAATA	2040
TCAGACAAC	AACGATGAAT	CAGCTAAAAA	ATATTTGGTG	GCGTGTCTTA	GTTACACAAA	2100
TTGAAAAAAT	TAGAACTTAA	TTATGACGTC	TATTTAGAGC	GATGGAATAT	GAGTATTAAG	2160

TTATTACTAA TATGGGTT

2178

## (2) INFORMATION FOR SEQ ID NO:200:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1528 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:200:

20	TCATTTTTTT GATTGAGCCA AAAGGAGTCC CCCCTGTGAG CGGAATATTT AATTTATTGA	60
	GCTATTTAAT ATTAAACGTA CGCTTAACCC CCCTANAGTG ATATCGTTTC TAGCGTTGTC	120
	ATTATCATTA AGCGAAACAT TTAAAGACA AATACACACT GTACGATCAC CAAACTGCAT	180
	GTCGAACAAT GTAACATTTG GATTCGATAT TTAATAATTGC TTGTGATGAT AAACCTTCTC	240
25	ATTTAGAAAA CGCTTCCACG TACATTCAAA AAAATAACTT TGTTAACCAT ATTGTAACAT	300
	TATTTTCATAT ATTTTGGGGC ATGAGAAATGA TTCTCACGCC CAGTAATTTA TTTATGCAAT	360
	TGTTTCATGTA GGTTCTTTGC GACGTTTTCA GGAATACCTA TATTTTFAAA ATCTTCAAGT	420
	GTAGCTCGTG CCTTCATTTT CTTGATTGAA CCGAATGAAC GCAATAATAA TGTTTAAACG	480
30	TTTGTTACCG ATACCATCTA TATCATCAAG TATTGATTTC AAGCCTGTCT TTTGACGTGT	540
	TTGTCTATGA AATGTGATTG CGAATCTGTG AACCTCATCT TGGATACGGA TGCAACAAAT	600
	AAAATGCCTG CCTATTTTTC TTCAGTGGAC AATTTCCTGC ACTAGCGCCA TAATAATAAT	660
	TCAGATGTTT GGTGTTTATC ATTTTCTCTG AAACCTGCAA CAGGGATATC AAGACCTAAT	720
35	TCGTTTGTGA GCACATCAAT AACCCCGTTC ATATGTCCTT TACCACCATC GACTATTATT	780
	AAATCAGGTA ATGGTAATCC TTCGTTTAAA ACGCGAGAAT ATCGTCGTCT TACTACTTCT	840
	CTCATTGATT TGTAATCATC TGGACCTTTA ACCGTTTGA TTTTATACTT TCTATAATTT	900
	TTCTTATCTG GTTTACCGTC GACAAATGTA ACCATTGCTG ACACTGGATC CACACCTTGA	960
40	ATATTAGAAT TATCGAATGC TTCAATTCTA ATTGGTGTTC GAATTCCCAT TTGTGTTCCA	1020
	AGTTCTTCAA TAGCTTTAAT CGTTCTTGAC TCATCACGTG ATATTAATTC AAATTTATTA	1080
	TTTAAGGATA CTTTAGCGTT ATGTGCAGCT AGGTCAACCA TATCTTTTTC GGGACCTCGC	1140
	GCGGGTTGAA CGATTTTAGT GTCCACAACA GATTGAATCA TTTCTTTATC CAAATTACGT	1200
45	GGTACATGAA CTTCTTTAGG TAAAATATGT TGGTTTAAGC TATAAAATTG TCCAATAAAT	1260
	GTATAAAATT CTTCTTCTTC TGTGTGGTGT AATGGAATCA TCGGNGTATC TCGCTTTATC	1320
	ATATTACCTT GTCGGATTAA AGAAAATTG GATACACATC CATCCTTTAT CAACACAATA	1380
	ACCAAAGACA TCACGAATCG GNTTATCTGA TGACATAATT TNGGGTGGGT GGGCAAGATT	1440
50	TGGATATTGT TGGATTAAAT CTCTATATTC TTNAGCCGCN CAAAATCAAG TGATTCACTT	1500
	GGAGGNCACA TNCGCTCTTC TAANCTTT	1528

## (2) INFORMATION FOR SEQ ID NO:201:

## (i) SEQUENCE CHARACTERISTICS:

# EP 0 841 394 A2

(A) LENGTH: 2933 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:201:

AGGTTGATAC	ATATGTCTTG	TAAATACATG	TGTCATTGCT	TGACCAATTG	CAGCAATCTG	60
GACCCAGGAT	NGGTGCCTAC	CCAGTTATCT	CCAAATTTCA	CATTCTGCG	TAATCTGGTG	120
CTACTTGAAA	TGACTTTTGT	ACCTTAATAG	CGGACTTCTG	TCATAAAATG	TGCATCCGGA	180
GTACGTGTTA	AAGGTACATT	AGAGCCCCAC	ATAATAATGT	ATGATGCGTT	ATACCAGTCA	240
CTTGATTGAG	GCACATCTGT	TTGCTCTCCC	CAAATTTGTG	GAGAGGCAGG	TGGTAAATCT	300
GCATACCACT	CATAAAAACT	AAGCATTTC	CCACCAAGCA	AATTGATGAA	TCGAGCACCT	360
GCTGCATAAC	TAATCATTGA	CATCGCTGGA	ATAGGTGTAA	ATCCTGCGAT	TCGATCTGGA	420
CCATATTTT	TEATTGTATA	CAGTAATTGT	GCTGCGATTA	TCTCTGTAAC	GTCTTTCCAA	480
TTTGAACGCA	CGTGCCCTCC	CATACCTCGG	GCTTGCTTAT	ATTGTTTGGC	TTTGTCTTCA	540
TTTTCACAA	TAGACGCCCA	TGCAGCAACG	CGATTACCAT	AGTTTCTTTC	TAATGCTTCA	600
GTCCATAAAT	CCCAGAGTTT	TCCACGAATA	TATGGATATT	TGATTCGAAG	CGGACTGTAT	660
TCATACCAAG	AGAATGACGC	ACCTCGTGGA	CATCCTCTCG	GTTTCATATTC	AGGCATATCC	720
GGACCACAAC	TTGGATAGTC	AGTTTGTGTA	TTTTCCTCAG	TAATCACACC	ATTTTTCACA	780
AATACTTTCC	AAGAACATGA	GCCTGTACAG	TTAACACCAT	GTGTGTGTTCT	TACTTCTTTA	840
TCGTGGCTCC	AACGTTCTCT	GTACATTTTT	TCCCATTTCT	TACTTTTACT	TTCTAGGATC	900
GACCAATTCC	CATTAAATTT	TCTGTTGGCT	TAAAGAATTC	AATCCAAATT	TTCCCATATT	960
TATATCTCTC	TACGTATAAA	AAATACGATG	TGTAGATGTC	GTGTNTTAA	ATACTTTAAA	1020
ATGCCCAAGA	CTATTGCTTT	AATTAGATTG	TACATTTTTT	CACAAATATA	AAATATTAGG	1080
GAATCACCTA	ATTACTTAAG	GAATTCCCTT	ATCAATAACG	GGATTTTCAAT	GAAATAATAC	1140
ACAATCATGT	ATGGTCGTGC	TTATTGCCAA	TCTAAATCGT	TCAAATTTGC	CACAACGACA	1200
AATAAGGCTT	CAACACGAAT	ATATTCTCTC	GGTTGAAACC	TTACTTATTC	ATTTATTTTT	1260
TATAAATTAG	TGACATAACA	CTGTATTAGC	ATCTGCACGA	TCGGTTGAAA	TATATGTTAC	1320
ATTTTCTTGC	TGCTTAATAA	ATGCATCATA	GTAATCATAT	TGCCGACCGG	AATGATATGT	1380
CCCATTCGAT	GTATCATTTG	GGTTTAGCAA	ACAGCCATAA	CCTTCGTCAT	ATAAATGTTC	1440
ACAGAGCATA	AGGGCGTCAT	GTCTAGAACC	ACTTACTACA	TAAAATTGCT	GAGTCGGATT	1500
AGAAGTTGTT	GAGACATTTT	CAGTATAAAC	CACTACTTCG	CCTATAATAC	ATATACCTGG	1560
TTTCGCCTCA	ATTGCGTAAT	GCTCCAGTTT	CCCAACAATA	TTACTTAAAC	GCCCCTTAAC	1620
AACAACTCG	TTAAACACG	ATGCTTGAAA	GACAATCGCT	ATCGGATAAT	CAATATCTGT	1680
GTAATGTTGT	ATCTGGGCGA	TAATTTTCCC	TAAACGTTTA	ACTCCCATAT	ATATCGCCAA	1740
CGTACCACCA	TTTACTAAAG	NGTTGACATC	TACGTCATTT	TCTTCTGAAT	CTNTNAAGTG	1800
ACCTGTAGAA	AATGTCNCAC	TCTTAGCAAC	AGCACGCNTC	GNTAATCCTG	TTTGCATCGT	1860
AGCAGCTGCA	GCACTCGCAA	GATGTCACAC	CAGGTACAAT	GTCANACGCA	ATAGTAAAAAG	1920
TTGTTTAGTG	TGTCAACTTC	TTCTTGAAC	CGNCCAAATA	TTGCTGGGTC	GCCACCTTTA	1980
ACCCTCACCA	ACCTTGTTAT	ATCGACGTGC	TGCTTCCACG	ATACATTCAT	TTATTTTTTC	2040
TTGCTGAATG	TGTTTTCAT	ACGGCTTTTT	ACCAACATCG	ATAATTTTCA	TAGTCAAATT	2100

EP 0 841 394 A2

	CGCATATTGT	AAAATTAACG	GATTCACTAA	TCGATCATAT	AGTATAACAT	CCGCTTCAAG	2160
	TATTAAACGC	TCAGCCTTTT	TAGTCAAATA	ATTCGGATTA	CCTGGTCCCG	CACCTATCAA	2220
5	GTAAACCTTG	CCATATTAT	CTACAGACAT	ATATATACGT	TCCCGTCTGT	AACTTCTACC	2280
	TCATAAACAT	CTACACAACC	TTCATCAGGT	TCTTGTACAA	TCCCAGTATT	TAAATCAATC	2340
	TTTTGATCAT	GAAGTGGACA	GAATACATAT	TCACCACTCA	CTGTTCCCTC	TGACAATGGT	2400
	CCTTGTTTAT	GCGGACAAAT	ATTGTGGAAT	CGCGTGANTT	GTACCACCTT	CAGATAAAAA	2460
	CCAATCCTAC	CTCTTTGCCT	TTGNCAATAA	CCTTTTTC	AATTAGGGGT	GTTAATTCAT	2520
10	CTATAGTTGT	CACTTTAATT	TTTTCTTTTG	TTTCCATGTA	TTACACCTTC	TCCACTTCAA	2580
	AAATTTTACG	TGCTTGCGCA	TTGCTAGNTA	TTGCTTCCCA	AGGTTTCAGCT	TCGACTGCTT	2640
	TTTTAGCATC	CATAATGCGT	TCAAATAGTT	CATTTTGTCT	TTCTGGGTCA	AGTAAGACTT	2700
	CTTTTACATT	TTCAAATCCA	AGTCTTCTTA	ACCATGGCGC	TGTTCTTTCA	GCATATATAC	2760
15	CTGTTTCGCG	ATATAGTGCA	TCAAAGCACC	ACATAATGTG	NTTACTTCAT	CTTCTGTTTC	2820
	TACAGTTGTT	AAAAATTCAG	CTTTTTCGAC	TTCTGTACCA	CCATTACCAC	CGATATAGAT	2880
	TTGGAATCCA	TTTTCAACTG	AGATAATACC	AAAATCTTTA	ACACCTGATT	CAG	2933

(2) INFORMATION FOR SEQ ID NO:202:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1866 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:202:

35	TAAAGTTCNT	TCCCCTGCTN	AATGTCCAAN	TTCCCCCAT	AGCAACTTTC	AAATATTTTT	60
	CATGACTTGC	TTTAGCCCAA	TCAAGTTCTT	TACCTGAAGG	TATATTAAAT	TGATTTGTTG	120
	AAAAGTTCCA	AAAATTCTGC	GCTTGGGTAA	GTCTTGTGG	GACAAATTTT	TGAAATCCTC	180
	CAACTTCTTA	AATATTTCTG	GTGATTTTGG	ATTAAAACTC	ACGTAATTTA	CGTAGTTTCT	240
40	CTTCTAATTC	ATGTTTTTTG	TTGACCTAAT	GTTTGTATTA	TTTGTGGTT	CGATGAAATG	300
	GCTTGCTGAT	TATCGGAAGC	ATGCTTTTTC	AAATTGTTAG	TTAAATTTTC	ATATCGCGTA	360
	ATTTGTTGAC	TTAATGATCT	GATATCTTCT	TCAAGCTCTG	ATTCTTTTAA	AGATATGCTA	420
	TCAACCTCAC	TCGTATAACG	TGACACAAAA	TTATCGCAAG	CTTGCTTCGT	TAAATCACTC	480
45	AATGTTTTCA	TACTTGTTGA	TAATGGAATT	AACACCGTAC	TAAAAAATTG	CTTAGCTGAC	540
	GTATACGCTT	TCCCTTTAAG	CGCATCATCA	TTGATAAATT	GAGTAATTGC	TTTTTCCAAC	600
	GCATCATAAT	TTGAATTCAT	TGTTTGACTC	AAATTCCCCA	CACTTGAAGC	TTGGTTTCGA	660
	GATCTGTCTA	AATACATGTC	AATACTCATC	GGCATGCTCC	TTTTTCAAAA	ATATATGATT	720
50	TTCAAACAT	TTAAATCAA	ATGCTTTTTC	CATCTACAAA	GTTGTAAAT	TTTAAACTC	780
	GGCGATGATT	ATTTCTTATG	TAAAGGAGTC	TAGATGCAGG	TAAATTGAGA	TAACATGTCG	840
	CCTTTTCT	TATTTTAGCA	TATGGATATA	ATGGTGTCTT	TGTATATTTC	CAATTAATCA	900
	ATAAAAATTA	TCTTTCAATA	TTTTAATTTT	ATTGCGACAA	CATCCTTAAC	ATTAAATATA	960
55	TTAATATCTC	AAAATATATT	CACTATTAAA	ATATGTCATC	AGTTGTAA	AGTATTTCTT	1020

EP 0 841 394 A2

	CATCATGCGA	AATATCAAAA	CGTATCTAAA	ATACGAATAG	TTTATCAATC	ACACAACATC	1080
	ATCATCCAAA	ATTTNATTGC	TCTCATNTNC	AATATTCATT	TCATATCTAT	CAGTTCATAC	1140
5	ATAATTAAAA	TTTCAATTGT	GCAATCTCAC	CGTTAATGCA	TCAACTTTCA	AACAATAAAT	1200
	CATCACAATA	ACCACACCTA	ATTCAACACT	TTTCAAACAT	AAGTATTGAC	ACATTGAGCA	1260
	AAATGATTTT	TAATTGTAAC	TAATACAGTT	ACAATTATGA	GGTGAGAAAC	ATTGAATTTA	1320
	GAATTTAACA	TTGCCGTGCA	TGTATTAGCT	TTTTTAACTA	AGCATCATTC	AGAAAAATTC	1380
10	AATAGTAGTT	CATTAGCAGA	ATTAAC TTGT	TTANATCCTG	TTCAATTACG	ACGCGTGACG	1440
	ACTCAACTTG	TCGATTTAAN	AATGATTGAC	ACAATACGAG	GTAAAGATGG	CGGTTATTTA	1500
	GCAAATGATC	AAAGTGCTGA	TGTCTCTCTA	GCAACATTAT	ATAAACATTT	TGTCTTAGAG	1560
	AAAGAACAAC	ACACACGTCT	ATTTACTTGG	CGACGAAGGC	AGTCACTGTC	AAATTGCTCG	1620
15	TAATATTGCA	ACTACCATGT	CACATTATCA	GCAAGACGAA	CAGAATATCA	TTATTAATTT	1680
	TTATAATGAA	AAAACAATCA	AAGATGTCAT	TGAAGACATT	CAAAAGGAGG	ATTTATGTCA	1740
	TGAAAACATA	TGATTTAATT	GTAATAGGAT	TTGGGNAAGC	TGGTAAAACT	TTAGCGAAAT	1800
	ATGCTGCATC	AACAGGTCAA	CACGTTTCGA	GGTATCGAAC	AATCTTCCAA	AATGTATGGN	1860
20	GGCACT						1866

(2) INFORMATION FOR SEQ ID NO:203:

25	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 2990 base pairs
	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: single
30	(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

35	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:203:
----	---

	ACAATGCCTC	CAAGAGGTAC	ANCCGATAAC	TGTCTCACAA	AACTAAATAA	ATTTAAC TCT	60
	GTGTCCTCTT	CCCATCCGTT	CAC TTTAATA	TCTTCTCCAT	AAGCATCAAC	GGAAAGATAA	120
40	AANCCGGACC	TGGAAATGTA	TGTGCCATCT	CTTTAACCCA	ATCAGTATCT	TGAATGCCTT	180
	TCGNTCCCAA	CCTATGCAAT	AATTAATCCC	TGCGGCAAAG	TAGTCCATGA	TTGGTGACT	240
	TCGTACGAAT	GCCACCACCT	ACTTCAATAT	CTTTGGTTGT	TAATCTCCTC	AATGACTTAA	300
	TATAATCAAA	CTCTCGGGCA	TACTGTGCCT	TAGCACCTAT	CAAGTCGACG	ATATGAATAC	360
45	GATTCACACA	TTCAAATTGA	CTATAGTAAG	CAATACTTCC	TTCAGCCGAA	CGCGACATTT	420
	TTTCTTCACT	ATCATATTTA	CCCTCTGTTA	ACCTCACACT	TGTTGACCCA	ATCAAATCAA	480
	TCGCTGGCCA	TAATTCAATC	ATTTATAAAT	CCCCCTTGTA	TTGCCTGACG	CAAAATTTGT	540
	AACCCATATG	TACCGCTTTT	TTCAGGATGG	AATTGAATAC	CAATATAATT	GTTAAATTGA	600
50	ACAATTGCCG	GAATATCAGC	CCCATACTGC	GCATATGCAA	TGACATTTTC	TGACATCGGC	660
	GCTTGGTAAG	AATGTACGAA	GTAAACATCT	TGATTTAACA	TAGGATGCTT	ACTCACTAAA	720
	TTATTCCAAC	CTAAGTGCGG	CACTGGGTAT	TCTGTTGGGA	TACGCGAAAT	ATTTCCTGGG	780
	ATAAACCTTA	ATCCAGATGC	ATCGCCTTCA	TCACTATGCT	CATACATTAA	TTGCATGCCT	840
55	AAACAAATAC	CAATCATCTT	CTTATCAGTA	TTCTTAGCCA	ATATTGCATT	GAGATTTAAT	900
	CGTTTTATCT	CTGACATCGC	ATCTTTAAAA	TGGCCGACAC	CGGGCAATAT	GATTGTTTCT	960

EP 0 841 394 A2

	GCTTGATCGA	TTATTTTGA	GGTATTGAG	ACAACCACCT	CATACCCTAA	ATGTTCAATA	1020
	GCGCGTTTTA	CATTACTAAT	ATCCCTAAT	CCATAATCAA	CGATGACAAT	CATTCAATCA	1080
5	CACCTTTTGA	TGACGGCACA	CGCTGATCAT	CAGTTGCAGT	TAGCGCTATG	CCTAATGCAC	1140
	GGGAAAACGC	TTTGAATATA	GCTTCAATTT	CATGGTGTGT	ATTACCTCCA	CGAATTAAAT	1200
	CAATATGCGT	TGTTAATCTT	GCATTGATTA	CGACCGCTCT	AAAAAATTCT	TCTACTAACT	1260
	CCGTATCAAA	CGTACCAACT	TTTTCTTTAC	TTAATGATGC	ATTGAATGAT	AGGTATGGGC	1320
10	GCCCCACTTAT	ATCCACAACG	ACACGTGCTA	ATGTTTCATC	CATTGGAATG	TACATCGTTC	1380
	CATAACGAAC	GAAATGCTTT	TTATCTTTAA	TCATTTCAAG	TAACAATTGG	CCAATGACAA	1440
	TGCCGATATC	TTCACTTACA	TGGTGATCAT	CTACGTCAAT	ATCACCTTGT	GCCTCAATGT	1500
	TTAATGACAG	ACCCTATGA	AATGTAAACA	ATGTTAACAT	ATGATTTAAA	AAGCCACAC	1560
15	CTGTATTAAAT	ATCGGATGGT	GACTGGTCAT	CTGATATTGA	AATATTTAGT	TGGGGTTTCA	1620
	GCTGTGTTTC	GGTGGTTGTG	AATAAATCAT	ATTGTGCGCT	CCANTCTTTA	ACAAATNTCTT	1680
	CCTAATTGCT	TTAACTGCGA	TGCTGTTGCA	ATTGAGTATC	TTACATAGCC	CTTCATCAAC	1740
	CGGCTCAACA	TAAAGCGAG	GTTTAAATCC	TTGTTCAAAT	ACGTATTGTC	CTAATTGGTG	1800
20	CGCTGCTGAA	CCTTTAGTAA	GTACAAAATT	GGCATTTGAT	GGGAATACTG	ACATTTTATC	1860
	TGCAACATGT	GTATCAAATA	TTTGTTTAA	CTGCTCAGCT	AACTGTCTGT	GCATCGTTAA	1920
	AAATACTCTT	GTCTCTTCTC	TATGTCTAAA	AATATAAGTC	GCAATATTTA	GCGTAAATAC	1980
	ATTAAATGGA	TATGGATGTT	CTATTTTGTG	AATATGCTTT	ATCGTTCCAG	CAGTACTAAT	2040
25	TAAGACACCT	AATCTTAAGC	CGGCAATTCC	AAACGCCTTT	GATAATGTAC	GCATTCTTAA	2100
	GATGTGTGGT	GCCAGTTCCA	CGTCATATGC	CGTACCATAA	TCTAAATATG	CTTCATCAAT	2160
	GACAAAGTAT	CCGTTTAAATG	CTTTCATCTT	ATCTGCAATA	GCTGTTAAAA	ATGCCGNATC	2220
	AAATTGCTTG	CCTGAAGGGT	TATGTGGATT	ACTCATAATA	AAAAATGATG	GNTGTNCTTC	2280
30	ATCGATTTTC	GTTAAATG	TTTCCAAATC	AAACGTNAAA	TCTGATCCNG	CATCTTACAA	2340
	ATGCAATTTC	ACGATTTACT	TGTGCCGCAT	ATGCTTGATA	CATAAAAAAA	TCAGGATTTA	2400
	GCGTTAATGC	CGGACCTTCT	GGCATGATCA	GCATTAACTT	TTGANTCANT	TCATCAGATC	2460
	CATTTCCATG	AATAANTTGT	TCAGGCGATA	ACCCGTAAAA	CTTAGCATNA	GCTTCCGTGA	2520
35	ATTGTTTATA	TGCTGCATCA	GGATATAAAT	TATATGGCGT	TGCACTAATA	ATAGACGTCA	2580
	TTGTTTPTTC	ATCCAACGGC	GTAACCTGGAC	TTTCATTTTT	ATCAATATAA	ATCATTTCAA	2640
	TTACACTCCC	CTAAGACTGA	CGTATTAAAA	TAGACTGCTG	GTGATTGTAT	AATGCTTCAA	2700
	CATGAGCAAT	ATGTTGTGCT	GAATCAGCAA	TTTGTTCAAA	CGTATCTTTT	GATAAATGGA	2760
40	TGACCGTGTT	CCGTGTTAAG	AAATCATTGA	CCGATAACCC	ATTGGTAAAT	CTAGCTGTTC	2820
	TATTTGTAGG	TAATACATGA	CTTGGACCTG	CAACGTAATC	TCCTATGACC	TCTGGCGAAT	2880
	AATGTCCAAT	AAACAATGCA	CCCACATATT	TCACTTNCNC	AATATATGGN	TGAGGATTTA	2940
45	CTGTGTGAAT	CGACGCATGT	TCAGGCGCGA	TCGTGCCCTT	CATGACATGG		2990

(2) INFORMATION FOR SEQ ID NO:204:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3386 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:204:

5	GCATCGGTTA AATGGCGTCA TTCCACTTAC GGGCATCTTN AGGCATTTTG CTTTTCAGT	60
	GGCTAATTTA GCTTCACCGT TTAAGGCTAT TTTTAGCATC AGTCACTTGA GTTAATGCAC	120
	GGAGTAATGT CATCTGGGTT AAGCGTTGGA TTTGTCGTTT GATTGATAAT ACCTTCCGCT	180
10	TGAGACACTG CATTTGTATA TGCTGTTTGC TTATCGACAT CAGCATCGTG GTAGTTCTCA	240
	CTTGCTTTCA CAGTATCTTT ATTTGCAATA CCTTGTTTCA AGTTACCCAT TGCTGTATTC	300
	ACATTTTGTG CATTTTGTG TAGTTGTGCC AACCTGTAGC TAAAGCTTGT GCTTGATTGA	360
	NTTGGTTTCG TAATGCATCA CGTTGTGGTG ATTTAAGTCA CGTNACGTAT CAAGATTTGC	420
15	TAAAGCATCT TGTTTCGCTT GCGCTAATTT TTCATCACCG TTTAATGCAT CTTTCGCTTG	480
	ATTCATAGTT GTAAGTGCCT TTGTTACTGT ATCAGGCGCC AATACAGGTG TTGGTGTGCC	540
	ATTTGCGATT TGTTGCGCTT GGTTAGCTGC ATGTTTATAA GCTTGTGTTT TGTCTTGATC	600
	CGCTTGTTGA AAGTCAACTG ATTTGTGGTAC TTGACTATTC GCTTGAATTT GTTGTGTCAA	660
20	CGTACCCATT GCATTATTTA ACGCATCAGC ATTTTGCTTA ATAGCATTAA CACCTGTAAC	720
	AAGTTCTGCA TCGGACACTT GGGTCTTTTA ATGCTGTTT TGTGGTTGA TTTAAGTTTG	780
	GTAAGTATGC AATGGCTGTA TTTGNATTGT CTTTAGCAAC TTGTAAGTTA TGGTTACCGT	840
	TTAAATCACC CTTAGCTCGA GCCACTTGTT GTAACGCTTG AGCCACTTGT TGTGGACCCA	900
25	CGCTTGCAAT TGGGTGNACC ACTAATGATT TGGTCCAGCA TCGCTACCG CATTTGAAAA	960
	ATGCATTTTT GCTTATCAGT ATCAGCATTT GCCAAAGTTA CCATTGCTT NAGGTGGTCT	1020
	TAATCTGCAA TACCTTGTNT CAATTGTGTC ATTGCTTTGA NNTAAATCTT TGTGCTGTTT	1080
	GCTTAATCGT GTTNACATTT GCAACAGTTT GCGCAGCATC GACTTGTGTT TTAAGGCAT	1140
30	CTTTGTGCGC TGGAATTTAA TCATTGTCAT TTGTAATAAC TTGTTTCGCT GCATCTTTTG	1200
	CACGTTGTAA CATTGTCATT ACCATTCAAT GCTTGTTCG GCTTCGTTCA CACGTTGCAT	1260
	CGCTTGTTTC CAACTTCAGT TTGAACCTGC ATTGCTGCCA TTTGCTTTAG ATAAATACC	1320
	TTCTGCATGT GAAACCGCAT TAGTATAATC ATCTTTCTTA GCTNGATCTG CATCAGTAAA	1380
35	GTAAATGCTA TCTAATGTCT CAGTTTATC TTGTAATGCA TTGTTTAAAGT TAGTCATAGC	1440
	ATTATTCAGA TTTTGAGCAT TTTGCTTAAAC ATTATTAACT GTTGCAATAT CTGGTGCTTG	1500
	TTCAACTTGA GTTGTAGCG CTTGTTTTTG AGCATCATTT AAATGATCTA GTGTACCAAG	1560
	TGTTGTCTTA GCATCTTGTT TGGCATCAGC TAATTTGTGA TTACCGTTAA GTGCTTGATT	1620
40	TGTGTCATTC ACTTTAGTTA ATGCTTGTTG AATTGTATTT GGATCCATTG TTGGATTGTT	1680
	TGTTTGTGTT AACAAATCTT CTGCTTGNGT TACTGCATTA TCATATGCTG TTTGCTTGTC	1740
	AGTATCAGCA TCGTGATAGN TCTCGCTACC TTTAATTTGT GCTTTATTCG CAATACCTTG	1800
	TTTCAATTGT GTCATCGCGT CATTAAGCGT TGTGCGCGT TGTGAATAG CATTGACGTT	1860
45	TGGTACAAGT GTTGCTTGCG TGATTGTGTT GCTGTATTCA TCACGTTGCG CTTGGTTTAA	1920
	GTTTGGAAT TGATCAATAG CATGTTGAAC ATTTTGTGTT GCCGTGCTA AGTTTGTATT	1980
	ACCATTTAAA TCATTTTGTG CTTGCGTAAC TNATCTAAC GCTGCAGGAA TTTGCTTAGG	2040
	TGTAACGACA ACATCAGGCG TACCACTAAT TAATGCTTCA GCTTTCGCTA CTGCTTGATT	2100
50	ATATGCATTT TTCTTGTCAG AATCTGCATT GACAAAGTTA CCATCAGCTT TTGTTGTGTT	2160
	TTTATCTGCA ATGCCTTGTT TTAATTGTGT CATCGCATTG TTTAACTCTT GCGCCGTTTG	2220
	TTTAACTTGG TTTACACCAG CTACAGTAGT TGCATTTGCT ACTTGTGTT TTAATGCATC	2280
	TTTCTGTGCT TGGTTAAGAT CATTAGAGTT ATTAATTAAT GCTGTTGCTT CGTCTTNTGC	2340
55	ATGTTGAACG TTGGCATTAC CATTTAATGC TTGTTNTGCT GCATTTACTT GTTGATTGTC	2400
	TTGNNCAACT TCAGTTTGTG TTGCATTACC ACCATTAGCT TGTGAAATG ATATTTCAG	2460

# EP 0 841 394 A2

5 CATTGTGAAC TGCAGTGTTA TATGCTTGTT GTTNGCCTTG ATCTNNATCA GNGAAGTTAC 2520  
 CTGNTGCTNG CGTAGCATCT ATATCACCTA TCGCTTGTC CAAGTTACCC ATCGCAGTAT 2580  
 TTAAGTTTTG CGCATTGTC TTAATTTGAT TTGCTTCATC ACCAGTATGC GCGCCATTAA 2640  
 TTTGATTAGT AACAGCTTGT TTTTGCGCAT TATTTAAGNT GTCTAATGAA CCTAAAGATT 2700  
 GCGTTGCTAG TTGCTTCGCT GCCTCTAAGT TTTTATTACC ATTTAAAGCA TTTTGTAGACG 2760  
 TGCTTACTTG TCCAGCAGCT TGATTGATAA CAGTCGGATC TAATGAAGGG TTTGTAGTTT 2820  
 10 GATCAATAAT ACCTTGTCGA GTTGTGACAG CATTATTGTA CGCATCCTTT TTATTCGGAC 2880  
 TTGCATCAGT ATAATTTTGG TTTTGTGTTG TTGTCGCATT ATCTGCAATA CTTTGACGTA 2940  
 ATTTGTCCAT TGCTGCATCA ACATTGTTTG CTTTTGTGTC ATTACCTTGT GCTTCTGCAA 3000  
 CAGTAGTCGA TTGTTGTACC AATTGTTTTA ATGCCTCTTT TTGTGCATTT GTTAAATGGC 3060  
 15 TTAAACCGTC AATTGCTGTA TTTGCGTGTT GTTTCGCTTT TTCAAGGTTT TGAGTACCAT 3120  
 TAAGCGCTGC TTTTGTAGTA TTCACTTGAT TCATAGCTGC TTCAACTGA TCTTTAGGCA 3180  
 CGTTCGTACC TGTAGATTAA TTTAAAATAT TTTTACGATT ACGAACCCTC TCATTGTATG 3240  
 CATTTTNCCT CTCTGGATCT TGCATCTTGC AAAGTNNNGG ACTGGCAAGT GTAGTGCTTT 3300  
 20 TATCATTCAA GCTATTTTTC AAGTTANCCA TCGCTGTTTC AAAATCCTGA GNGGGNTTGA 3360  
 TTGTTGGNG TTNACTTCAN CTACAG 3386

## (2) INFORMATION FOR SEQ ID NO:205:

25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 761 base pairs  
 (B) TYPE: nucleic acid  
 30 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:205:

40 AAAATACAAG CGTCTAAATA TTAACAATCT NCCTGTGACG CTATATGGCA TATCTTGCTT 60  
 TTTAAGCATC TGTTGTATCT CTGCCATCGA GCGCATTTGGC CAGCTTCACA ATAATTATAA 120  
 TGAATTGTTT TTATCGATGG AGACACTAAT TGTGTCATTG GGTCACCACC ANAACCATAT 180  
 ATTTGATGTG GTTTCATAAC ATCATTTTTTA TCAGAATAAT ATTTATAGGC AGCTAATGCA 240  
 ATCGTATCAG TTGCTCCAAC ANCCGCATCT ACTTGCTCCA CATTTTCCAA AACATTTGCA 300  
 45 ACATCTTTTT GTGCTCCAC ATAAGTAAAA TTTGTTTCAT GTATATTAGG TTTAATTTGG 360  
 TATTTAGCTA ACTGGTCCAA GTAAACCCAC GTTTTCTATG AATACCACCT GCAATATCTT 420  
 TTTCACTTAC ACTAAACACT TCAACTTGTT GATATCCCTG TTGACCAATC CATTCGCCTA 480  
 TAATTTGACC TGCTTTATAA TCATCATGCA CAATACTATG AAGTTGTTCA TGTGTTGAC 540  
 50 CAACAATAAC GATTGGTACA TTCATTTTAT TAATGACTTC AATATGTCTC TCTGTTATGT 600  
 CTGTAGCCAT TAAAACAATA CCATCTACTT TACTGCGTGC TAATGTTTCA AGCGCTTGTA 660  
 TTTCTGCTTC GATATTTAAA CCTGTGTAAT TTTAAATTAA TTGTGATTCA TATTTTGGC 720  
 ATTGTTTTGC CAATCCTTTG ATTGTTTCAT CTACTGCATA T 761

55 (2) INFORMATION FOR SEQ ID NO:206:



## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2334 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:206:

```

5  CCACCCANCT GATTATAATG TTTTAGCANG AGCTAGACTT GGTGGTTAC CATCATATCC      60
   ACAATTTAAT AAAAATAGTT TGTGTGTTGC AGAAGAAGCT AAAGATGAAG GCATTGAGTC      120
   GAATGAGGCA ATTTTAAAAC GAGCGATAAA TGGAAAGTTAA GTCAAAACAA ACGCAATTTG      180
10  CGATAGAAGA TCCGGATTTG AAAAAGAATC ATCCGGAAAT CACTGTTTAT ATGGCGCTCA      240
   AATCTAATCT CAAGTCTGTC AAAAGGTCAA GAATACTTTA TGAAGCATTT ACTTGGCACA      300
   AAATCAGGGT TATTAGCTAC ACCAAATGAA GATGAAAAGC CAGAAGAAAT TACGTGGCGT      360
   GAGGAAACAA CAGGGAATTT AGATTTAGTC GTTCTCTTAG ATTTCAGAAT GACAGCAACA      420
15  CCTTTATATT CTGACATTGT TTTGCCAGCA GCGACTTGGT ATGAGAAGCA TGATTTGTCA      480
   TCTACAGATA TGCATCCATA TGTACATCCT TTTAATCCAG CTATTGATCC ATTATGGGAA      540
   TCGCGTTTCA ACTGGGATAT TTATAAACG TTGGCAAAAG CATTTTCAGA AATGGCAAAA      600
   GACTATTTAC CTGGAACGTT TAAAGATGTT GTGACAACCT CACTTAGTCA TGATACAAAG      660
20  CAAGAAATTT CAACACCATA CGGCGTAGTG AAAGATTGGT CGAAGGGTGA AATTGAAGCG      720
   GTACCTGGAC GTACAATGCC TAACTTTGCA ATTGTAGAAC GCGACTACAC TAAAATTTAC      780
   GACAAATATG TCACGCTTGG TCCTGTACTT GAAAAAGGGA AAGTTGGAGC ACATGGTGTA      840
   AGTTTCGGTG TCAGTGAACA ATATGAAGAA TAAAAAGTA TGTTAGGTAC GTGGAGTGAT      900
25  ACAAATGATG ATTCTGTGAG AGCGAATCGT CCGCGTATTG ATACAGCACG TAATGTAGCA      960
   GATGCAATAC TAAGTATTTT ATCTGTACG AATGGTAAAT TATCACAAAA ATCATATGAA      1020
   GATCTTGAAG AACAAACTGG AATGCCGTTA AAAGATATTT CTAGCGAACG TGCTGCTGAG      1080
   AAAATTCGTT TTTAAATATA ACTTCACAAC CACGAGAAGT AATACCGACA GCAGTATTCC      1140
30  CAGGTTCAAA TAAACAAGGT CGACGATATT CACCATTTAC AACGAATATA GAACGCTCTAG      1200
   TACCTTTTAG AACATTAACA GGACGTCAAA GTTATTATGT GGATCACGAA GTTTTCCAAC      1260
   AATTTGGGGA GAGCTTACCA GTATATAAAC CGACATTGCC GCCAATGGTA TTTGGGAATA      1320
   GAGATAAGAA AATTAANGGT GGTACAGATG CTTTGGTACT GCGTTATTTA ACGCCTCATG      1380
35  GANAATGGAA TATACACTCA ATGTATCAAG ATAATAAGCA TATGTTGACA CTATTTAGAG      1440
   GTGTCCACCG GTTTGGATAT CANATGAAGA TGCTGNAAAA CACGATATCC AAGATAATGA      1500
   TTGGCTAGAA GTGTATANCC GTAATGGTGT TGTAAACGGCA AGAGCAGTTA TTTCGCATCG      1560
   TATGCCTAAA GGTACAATGT TTATGTATCA TGCACAAGAT AAACATATTC AAACGCCTGG      1620
40  GTCAGAAATT ACAGATACAC GTGGTGGTTC ACACAACGCG CCGACTAGAA TCCATTTGAA      1680
   ACCAACACAA CTAGTCGGAG GATACGCACA AATTAGTTAT CACTTTAATT ATTATGGACC      1740
   AATTGGGAAC CAAAGGGATT TATATGTAGC AGTTAGAAAAG ATGAAGGAGG TTAATTGGCT      1800
   TGAAGATTAA AGCGCAAGTT GCGATGGTAT TAAATTTAGA TAAATGCATA GGATGCCATA      1860
45  CGTGTAGTGT GACATGTAAA AACACTTGGA CAAATCGTCC AGGTGCTGAG TAACATGTGG      1920
   TTCAATAACG TAGAAACGAA GCCAGGTGTA GGGTATCCGA AACGTTGGGA AGACCAAGAA      1980

```

EP 0 841 394 A2

CACTACAAAG GTGGTTGGGT ACTAAANTCG TAAAGGGAAA CTTGAATTAA AATCTGGAAG 2040  
TAGAATTTCA CAAATTGCTT TAGGTAAAAT TTTTATAAC CCAGATATNC CATTAAATAA 2100  
5 AGATTATTAT GANCCATGGA NCTATAATTA TGAACATTTA ACAACTGCCA AATCAGGGAA 2160  
GCATTGCGCA GTTGCTAGAG CGTATTCAGA AATTACAGGG GATAACATTG AAATTGAATG 2220  
GGGACCTAAC TGGGAAGATG ACTTAGCAGG TGGTCATGTT ACAGGCCCAA AAGATCCTAA 2280  
CATACACAAA ATAGAAGAAG AGATTAAATT CCAATTTGAC GAAACTTTTA TGAG 2334

(2) INFORMATION FOR SEQ ID NO:207:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2698 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:207:

GGCACGAGAC TAATACCTGA AATTAATCCA CCACCACCAA TAGCGGCAAA TAAATAATCA 60  
AAGTTAACGG AATTGTCAGT CTTTGATTGT TCTAGCATTT CTTTAGCAAG CGTACCTTGT 120  
CCAGAAATTG TATGAACATT ATTGAATGGA TCTATAAAGT TCATTTGATG TTCACTTGTA 180  
30 TAAGTTAAAG CTTGAGCTAA ACAGNGATCA AATGTATCAC CAGTGAGTAC AACTTCAACG 240  
TTACTATTTC CAAAGAACTT TACTTGATT ACCTTTTGTN AAGGTGTAGT GACTGGCATA 300  
AAGATAACAG CGTTTANATT AAGTTTTTTA GCTGTATAGG CAACACCTTG AGCATGATTA 360  
CCTGCACTCG CACATGTAAT ACCTTTACTT TTAGCTTCAT CTGATAAAAC AGAAATAGCG 420  
TTGTAAGCAC CTCTTAATTT AAAAGAACGT ACCCATTTGTA AATCTTCTCG TTTTAAATAA 480  
35 ACTTTACAAT CATACTTTTG AGATAAGTAA TGGTCTAATT GTAAAGGTGT TTCTTTGACA 540  
ATATCTTTAA GTCTTAAAAA TGCTTCATCG ATATCTTNCG TAGAACTGN TGTTTTGACT 600  
GGCATAATAT TCAACTCCCT TAAAGTGATT TNNCATATTN NTCTATTAAT GATTCATATT 660  
GTAGGGTGAT TGCAATTGTC ATCTAAGCCA TTTACCAAGT TTATTTTTC AAAGTTTCAT 720  
40 CAATTTCAAA ATGGAAACGC TTGCTGGTG ATGACACAGT TTGATTGGT AAATCTATTT 780  
CAATTTCCCT ATATTGTGCA AGATGTTTAC GAGCACTTTT TTCTAAANCG ATAGGCAACA 840  
TCGCATTTT AGTGCAATTC ATATAAAATA TGCTACTGAA ACTTCCTGCA ATAATAATAT 900  
45 GNAAACCATA GTCCTTAAGA GCCCAAGCAG CATGTTTACG ACTTGAACCA CATCCAAAAT 960  
TATCTCCAGT AATTAAAAATA GAAGCCCTT TATATTGTGG TTTGTTAGGA TTGAAATCAG 1020  
GATTATCTGA ACCATCAGGT AAGTACCGCC ATTCATCAA AGCAAATGGA CCAAAGCCAC 1080  
TTTTTGAAAT ACGCTTTAAG TGTACCTTAG GAATGATTG GTCTGTATCG ATATTGTCGT 1140  
50 TGAAGAGAGG GACTATTTTA CCTTTATATG TTGTAATAGG TTTGATTGCT GCCATTTAAA 1200  
CAACCACCTT TCTTACGTCC ACAAATTTAC CATGAATAGC TGCTGCTGCT GCCATAGCAG 1260  
GGGATACTAA ATGTGTTCTT GCACCTTTGC CTTGTCGTCC TTCAAAGTTT CGATTACTTG 1320  
TAGATGCACA ATGTACGCCC TCAGGTACTT GGTGAGGAT CATGCCTAAA CACATTGAAC 1380  
ATCCTGGTTC ACGCCATTCA AATCCTGCAT TTTTAAAGAT AGTATCTAGA CCTAATTTT 1440  
55 CTGCTTCTTT TTTTACTGTA CGAGAACCTG GTACGACAAT AGCTGTAATA TTTGGATGAA 1500

EP 0 841 394 A2

	CTTTATTTCC	TTTAACAATA	TGACTAGCTT	CAATCAAATC	TGATAGTCTA	GCATTGTAC	1560
	ATGAACCGAG	AAAAACATAC	CCTAAGTCGA	TGTCCTCAGC	TTTTTGACCT	GGTTCTAACC	1620
5	CCATATAATC	ATACGCACGT	TGATCGTTGA	TATCATTGAT	TTCAGGGAAT	GGTTCAGTGA	1680
	AATTAACACC	CATTTTCAGGA	TTAGTTCCCC	ATGTCACCTG	TGGTTCTAAT	GTTGAAACAT	1740
	CAAGTTCAAT	TACACGATCA	AATATCGCGT	CGTGCCATCA	GAATATAGCT	CACGCCACTT	1800
	ATCAACTGAT	TTAGCGAAGT	TATCGGCAAA	TGGTCTCCCT	TTAACATATT	CAAATGTTAT	1860
10	ATCATCAGGT	TGGATTATGC	CGTATTTGGC	ACCACCTTCG	ATAGCCATGT	TACAAATAGT	1920
	CATTCGACCG	TCCATTGAAA	GGTTTTTAAT	TGTTTTCGCCA	GTAAATTCCA	AAGCATAGCC	1980
	TGTACCAAAA	TCAACACCAT	ACGTTTTAAT	TAAATGCAGA	ATAATGTCTT	TAGCATAGAC	2040
	GCCTGTTGGT	AAGGTACCAT	TAATATCGAT	TTTTAAGTTT	TTGGGTTTGT	TTTGCCATAG	2100
15	CGTTTGAGTT	GCGAAAACAT	GTTCAACTTC	ACTTGTTCCA	ATTCCAAATG	CAATAGCACC	2160
	AAATGCTCCA	TGTGTTGCTG	TGTGAGAGTC	ACCACAAACG	ATTGTTTTGC	CGGGCTGTGT	2220
	AAGTCTGTG	TCAGGTCCTA	CCATGTGAAC	AATACCTTGT	TCATCAGAGC	CCATATCAAA	2280
	AATATGCACC	CCAAAATCTA	TGGCGTTTTT	TTGTAATGTT	GTGATTTGTT	TGTTTGCAAT	2340
20	TTCACTTTTA	ATATTGAAAA	TATCAATAGT	AGGAACATTG	TGATCGAGTG	TTGCAAAATG	2400
	TAAATCTGGG	CGTCTTAATT	TTCTGTTTTG	AAGCCTAAGT	CCTTCAAATG	CTTGAGGAGA	2460
	AGTAACTTCA	TGTATAAGGT	GTAAATCAAT	GTATAATAGT	TGCGGGTCGC	CCNATTTCCC	2520
	GNATAACACA	TGTCTGTTCC	ACACCTTGTC	AAATAATGTT	TGACCCAGTT	ACATCTCCCC	2580
25	CTNAGTGATT	CAATTTTTGA	GATAGCATTT	CGAAAAATATC	AGTAGTATTC	AATTTGCCGC	2640
	CTAAATCTGC	TGTCGTTTGC	CCATGTTCAA	TCATGTTATA	AATATGTTGT	TCTAATTC	2698

(2) INFORMATION FOR SEQ ID NO:208:

30	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 706 base pairs
	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: single
35	(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:208:
----	---

	CTAAGANGAT	ANTTTAAGCT	TGTAACCCAA	CAGTTCGCAG	GTAAATCCA	ATCCAGTTTG	60
	ACCAAAGAAA	TACCGACCCA	AACCGCCTTT	GNACCGTAGT	TGGATAATTA	ATGGGGACCG	120
45	CCCAAAGATA	ATAAGCAGTT	AGCCGNNNAG	GTTCCGGGAA	TAGGNCCGTT	TAAATTNAAA	180
	ACCCAGCTGC	CTTTTGCTTC	TTGAAATACA	TTGGGGAATC	TGGATTAGAT	GGTCCGAAAC	240
	TACCGTCGAC	ACCATCACGT	CAAGCAGTTG	CAATNCCAGC	ATCAGATGTG	ATTATAAACA	300
	AAGTGATGAC	GAGTGATACA	CCTGTAACAA	TTGTAGCGAC	AGGTCCTCTT	ACGAATGTAG	360
50	CAACGGCATT	GATTCGTGAT	CCAAGAATCG	CTGAGCATAT	TGAATCTATT	ACTTTGATGG	420
	GTGGTGGTAC	ATTTGGAAAT	TGGACCGCCT	ACAGCAGAAT	TCTATCAATA	CCTACTTCTT	480
	TCTTGACTAA	ATCANACTGT	GGCTTCGTTA	ACATGCCACT	TGGTGTGATA	TAAAACTAT	540
	TTTTAAAGTA	ATAGCTTACA	GAGTGGTTCA	AATGTTTCAGC	GAATAAAGCT	TCATCCATAC	600
55	GTTCCTAAGAA	GAATGGGATA	AACTCACCCC	AATGTCCAAT	AATCATATTT	AACTTTGGAT	660

AACGATCAAN AATGCCAGAT AATACTAGAT GTATTGCATR RDCTNS

706

## (2) INFORMATION FOR SEQ ID NO:209:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 634 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:209:

CTTGTTGGAT CAGATCCAAC ACTACCTGAT GGATCNCAT CNAGATTGAG CACGCCACCA 60  
 TATCAGATGG AATATCAACC TTGTCATGGG ACAGACTGAC ACCTGAATGA NGGGGTGTCT 120  
 GGNCANNAC ATCATCACCA TTAATGGGAC AGCGAGTGGC ACATNTCCAT GTAGTAAATG 180  
 GCACTTATTA CTTACACGGN CATATCGTGN CAGGTTGGCA AGGTGTGAAA AAGACATGTG 240  
 ATACAGCGGA AGAGCTTGAC ACATATATAA AGCANAGTGA TGTGGTATAT GAGGAACAGA 300  
 AGCATCTANC TTTATTTTAA GNGGGCGGAA ACAATGAAAC TCAANGTTAA AAGAGANATG 360  
 AGATTAGATG AATTAATTAA NTGGGCGAGA GAAAATCCGG ATCTATCACA NGGAAAAATA 420  
 TTTTTCNCAA CAAGGATTTA GTGATGGATT CCGTCCGNTT TCATCCANAT ACAAATAAGT 480  
 GTCTCGACGN CAAGTTTTAT TCCAATTGAT ATCCCCTTCA TAGTTGATAT TGAAAAAGAA 540  
 GTAACGGAAG AGACTAAGGT TGATAGGTTG ATTGAATTAT TCGAGATTCA AGAAGGAGAC 600  
 TATAACTCTA CACTATATGA GAACACGRRD CTNS 634

## (2) INFORMATION FOR SEQ ID NO:210:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1157 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:210:

GGCACGAGCG CCCATTTAAT TAATTCATCT AATCTCATTT CTTTTTTAAC TTTGATTTTC 60  
 ATTGTTATAT CTCCTCTTGA ACAGTAAATT TATCGTTAAT TGATACGTAT CCAGTCACAT 120  
 TACATAAGAT GCTATCAACA TCAAAAGTCA CACAACAGTT GCGTTCAACA TCATTGAAAT 180  
 AGAATCTTTT ATTACCTGAT AACTTGCGGG TTATCACCAA GCCCATTGAA TAAATTCAGC 240  
 TAAATCTCAT TTCTTTTCA ACTTTGATTT TCATTGTTTC CGCCCTTTTA AAATAAAGTT 300  
 AGTTGCTTCT GTTCCTCATA TTCCAGAATC ACTTTGCTTT ATATATGTTT CAAGCTCTTC 360

EP 0 841 394 A2

	CGCTGTATCA AATGTCTTTT TCACACCTTG CCAACCTGGC ACGATATGAC CGTGAAAGTA	420
	ATAAGTGCCA TTTACTACAT GGATATGTGC CACTCGTTCG TTATCCTGAT ACAGATATCT	480
5	CTTAGAGCCG AAAAAATGTT TTAAGTATTC TTTACGTCCG CTATCTGTCA TGGTCATCAC	540
	TCCCGCAAGT CAAATACTCT ATCGACGTAA AACTTCGCCT TTGCTAAATC CTCATGACCA	600
	TTCTTTAACG GTGCTCTAGA CAAGTATTTA ATTGCATTAC CTATTGCGAA TGCTAATTGT	660
	GGTGGGTACT GTGCCGTAAC TTGTTCAATA AAATCTATAA TTTCAATGTT CGCCGTATGT	720
10	TGTTAATGCG CAGGTTGCTT AACGTTGTCT TGCCTTTTGT TCATATCTAC TTTTCTGTTA	780
	CTGATTATGC TCATTATGTT TCACTCCATT TCTTGAACAT TTGGTTATAG TTGACATCGA	840
	ACCAGTACGA TCACGTTGAA TGTTTTTGAG TACATCAAAT AATGTTGCTC CTCTCTTAGC	900
	TCAGCTTCTT TACGTCGTAG CCTAGCCATT TCACGCTCTC GCTCCAAAGC TTTTGTATT	960
15	TGTATTTCTC TATAGTCGTT TAGCTTCATG CCGAAAGGTG CATCAATTGC TTCCGACATC	1020
	TCCCAACCCT TCGCAACTCT GTTCTAACT ATTCGGGCG TGAGTCCTTT CTTTTTCATC	1080
	TGCTCATTTT CATATTCAGT GTATTTAGAA GGGGGTTTTT CTTGTGGTGG CGCAATAAGC	1140
	GCATCGCCCG TTAGCCC	1157

(2) INFORMATION FOR SEQ ID NO:211:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 725 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:211:

35	GGCAGGAGCT CGTGCCGATT CGGCACGAGC .TTCGGCTGTA TTGGTACCTA TCAATACAGA	60
	CATTGCCGCT TTTTCTAAAT CTTCAATACT ACCTTGAGTC ACCCGCCGTG TCCAACCGTT	120
	CTAAATAAGC CTTTGTGTC AGGTGCTTGN CAATATCTTT GNACAAGCTA TCAACGTTAT	180
40	TTGTGCCAAT AACTTGCTCT GCTAATTTTT GCANTAAANA CGATTCTTCA TTCGTCGCTT	240
	TAGAAGAAGA AATGAATGAT AGTGCATCTG GGCCATACTT TTCTTTAATA GATGTAAAAT	300
	TATCTGCAAT GACATTTAAA GCTTCATCCC ATTCTACTTC ATGGAAGTCA CCATTTTTC	360
	TTACTAGTGG TTTAGTTAAG CGTTGATCTG AATTAATATG TCCCATGAA AACTTACCTT	420
45	TAACACAAGT CGCAATTTTA TTTGCTGGAG AATCATGTGA TGGTTGTAAT TTTAAAATTT	480
	CTCTATCTTT AGTCCAAACT TCAAATGAAC AACCACACC ACAATAAGTA CACACTGTTT	540
	TAGTTTCTTT AATACGCTCT TAACGCAGTT CTGCTTCTGA ATCTGAGATT GCAAATAGTG	600
	GACCATAACC AGGTTCTGCT TTTTGTAGTA AATCAATCAT TGCTGCTAAT GAACCAGGTT	660
50	CCGTATCAGT CATATAACCC GCATTACCTT CCAGATTCAC TTCCATCATG GCATTACATG	720
	GACAG	725

(2) INFORMATION FOR SEQ ID NO:212:

(i) SEQUENCE CHARACTERISTICS:

# EP 0 841 394 A2

(A) LENGTH: 1676 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:212:

AAAAATCCAAG	GCACAAAGTT	CTAAATACGG	AGTCAAAAGA	ATATTGAACA	AGCACATTTT	60
CAACAATACT	TTGAAAATTT	ATTATTTCAT	AAGTTCATCA	TTGATCCAAA	GTAATATTGC	120
CCTTGCCGAT	GTTCACTCTA	TTGACTATTT	ATTTACAGCC	ATTAGTTAGA	TATATTTTTC	180
AGCGAATTGT	CATGGCTGTG	ATTGTCATCA	TTGGTGTTAT	TGTCAGTGTG	TTTACCATTTC	240
TGTATTTTTC	ACCGCTTGAT	GCGGCTTATA	GCATACTGGG	ACACAATGCA	ACANAGGCAC	300
AGATACATCA	ATTCAATGTA	TTACATCATC	TTAACGAACC	TTATTTTATT	CAATTGTGGG	360
ATACCATCAA	GGGTGTTTTT	ACCTTTGACT	TAGGTACGAC	TTACAAAGGG	AATGAGGTTG	420
TGACTAAAGC	AGTTGGCGAA	AGAAATCCAA	TTACAATAAT	TGTCGCAGTA	TTAGCGCTAA	480
TTGTGGCATT	AATTATTGCA	ATACCAATTG	GTATTATCAG	TGCGATGAAG	CGAAATAGTT	540
GGCTTGATAT	CACGTTAATG	ATAATTGCAT	TAATTGGTTT	ATCTATTCCA	AGTTTCTGGC	600
AAGGGCTATT	ATTCATTTTA	GCGNTCTCAT	TGAAATTGGA	TATTTTGCCA	CCATCTTATA	660
TGCCAGAACA	TCCAATATCG	GTGGATTTTA	CCTGTACTTG	TCATTGGAAC	AAGTATTGCT	720
GCTTCTATCA	CGCGTATTGA	CAAGGTCTTC	TTGTACTTGA	AGGTAATGCG	CAGCGATTAT	780
GTTTTNACTG	CTTATGCAAA	AGGATTATCG	ACGACACAAG	TTGGTTATTA	AACATATTGT	840
GGAAAAATGC	CATTATTCCA	ATTGTACGGT	AAGTGGGTCT	TCCTNGGTGG	CAGAGTTACT	900
AGGCGGTTC	GCAGTGACGG	AACAAGTATT	TAACATTAAT	GGTATCGGTC	GTTATATCGT	960
CCAAAAACAA	CTAATACCTG	ATATCCCAGC	AGTCATGGGT	GGGGTCGTAT	ATATATATCA	1020
ATTGTAATAT	CTTTAGCAAA	CTTAATTATT	GATATATTTT	ATGCTTTAAT	CGATCCAAAA	1080
TTACGTAGTG	AAATTAACGA	AAGGAAGTGA	GGCATATGGT	AAACTTACA	ACAAAGATAG	1140
CTTCCTTAAA	ACTATTTCGA	AGTTATGCTA	TAGCAACTTA	TATTTTAGTT	ATGTTAACGA	1200
GTGCATTAAA	TCTTTTAAA	GGTTATGTGG	CCGATACGTT	CTATATTGCT	GAAACATTGC	1260
TAATCATTTT	AACCATCATT	TTAATTATCA	TTTAAACAAC	GGAACAAACA	TGGAAGCATC	1320
ATGACCTATG	GCGACGTATC	GTGGAAGTGT	TGTTATTGTT	GATGGCATT	ACAGGCAACG	1380
TATTTACATT	ATTAATGTGT	GTAAGTATTA	GACGTTACCA	ACGTACATCG	CAAATACATA	1440
GTTATAACGG	GGTGGGAATC	CGNTTATACG	AAAACTACT	AGACATCCGT	ATTTGCGATT	1500
ATCGGGGTTA	CTTATTTTAG	TCTACATGCT	GACATTATCA	ATNGTGTCAC	AATTTAAATT	1560
TGATACGACA	ATGGCTTACT	AAAAATCCAG	TCCAATGAAC	TNGTACAATG	GACCGAGTCT	1620
AGCCTATCCG	TTTGGTACTG	ATGATTTCGG	TAGAGACTTA	TTTACACGCT	CGTGCC	1676

(2) INFORMATION FOR SEQ ID NO:213:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1978 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

# EP 0 841 394 A2

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:213:

10	CTAATCGTAT GAACGAGTAC CCAGAGGTTT GAAATTAAAT ATTGATTCTA TNAATGATTTC	60
	CTTAGTGTGCG CATAACGGTG CAAGAGCACG ATACTTAGGA TCAATAAAAC CTTCTGTCAT	120
	CATATGGTCA ATCATTGNTT GTAGTGGNTT GAAAANGCCA TTAATATTAT AAATGGCAAT	180
	AGGCTTGTC A TGATACCTA TTTGAGCCCA ACTAATACAT TTCGAAAAAT TCTTCTAGTG	240
15	AACCTTGCGC CACCAGGAGC CATGACAAAT GCATCTGCAA GTTCTGCCAT TTTATTTTTA	300
	CGTTCATGCA TAGAATCCAA CTAAAATTAA TTCAGTTAAA CGTTGGCTTG TGATTTTCATG	360
	TTCATCTAAC ATTTTAGGCA TGACGCCAAT AGCTTTGCCG CCATGATTTA ATACACCATC	420
	TTGAATGGCA CCCATAATGC CAATTGACCC TGCACCAAAT ACTAATTCAT AACCTTGTTTC	480
20	AGCAAAATAT TTACCTAAAT CGTATGCTTT TTGTACATAT GAAGGGTCAT GACCTTTGCT	540
	TGCACCACAA TAAACTGCGA TTCGTTTCAT GTTAATCCAG CTCCTTAATT CGATGAATGA	600
	CTTTTAATAG TGATTGTTCA AACACTTTTT GATCTTGCTT TGTAAGAGGT GGGGGACCTT	660
	TGTGGCGACC ACCTTGTTTT CTAATTTGTG CATTCATATA TCGTTTATCT AATAGTTGTT	720
25	GAATATTTTT GGAATTGTAT ATCTTCCCAT TATGATGCAT GACAATTAAG ACTTTGTCGA	780
	CTAATAAACT TGCAGTCCA TAATCTTGAG TGACTACGAT ATCATCCTTC GTTGATAATT	840
	GAACAATTTT GTAATCAACT GCATCTGGTC CATCATCAAC ATATAATGTT GATACATGTG	900
	GAGGATATAA TTGGTTCGAA AAATGGCTGA AGCTCCGAAT AATTGTCACA AAAATGCCTG	960
30	TCTCAGTTGT TAAATCTATA ATAGAATCNA CAACAGGACA AGCATCTCCA TCAATAATAA	1020
	TATGTGTCAC AATTATGCCT CTGTATTGTT TTCTTTATTT TGTGAGAGG CGCTTTTGCG	1080
	AACATAATCT TTATATTTTT TAAATGACTT GATGCGTGCT TTATCAGCTT CTGTTGGCG	1140
	TTTGTTCTTC TTTGTGTCGT TTTTCAATAT TTTTGTGTA CTTTTTATTC ATTTTAGCGA	1200
35	TTTCTTTGCG ATTTTTTTCA GCTAGTTTAT CGCCTTTTCT CTCAGTTTTC TCATCTAATT	1260
	TATTAGGTGT TAAGCCTGCT TTTTCTTCGT ATTTTGTGTA TTTTTCATA TCTTTAATAC	1320
	GTTGTATTTC ATTCTTTTCG CGGGCTTTTT GCTCTTCTTT ATGACGCTTT TCGATATTTT	1380
	TTTGAAGTAT TTTATTCATT TTATCAGCGT CTTTACGATT TTGTTTAGCT AATTTTTCGC	1440
40	CTTTTTTCTC AATATAGGCA GGATCATGTT CTCTAGCAA CTTTTTAAGT TCACGTTTAT	1500
	TTTCAAAATC TTGTTTTTTA TCGCCGACAT ATTCTTTAAC ATCACTCGCT GTGTTACTGA	1560
	TTGCTGCAGA TGTTTTTGAA GCAACTTTAC TTGGTAGCAT CTGTAACTTT TGGGACGTCC	1620
	GGATGTTGTT TGATACGTTT ACGTTCAACA ATTAACGGTA CCAATACAAT TGGTAATACA	1680
45	TTAATCATAA ATTTGATGAC TTTTCTCTTA TCCATAGATC TTGCCTCCAT AATTACTTTA	1740
	TTAATTTTAC ATACCTATG ATACATCAAT ATAAACGATG ATAGTAGTGA NTCACTATTA	1800
	AGTATTTTCA ATGTTTTTTA NAAGAAGACA ATAAAACTG CCAATCACGT GGTTCCTTAA	1860
	TTGACAGTCT ATATTTTANA CAGGAAATTA AATACCTTTA CCAATGCCAA TCCGAAGTAA	1920
50	AGTATAGCAA TANAGATTAC TAATACANTT CTGCTAAATG GCANATGGAA TTAGTCTG	1978

(2) INFORMATION FOR SEQ ID NO:214:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2558 base pairs

EP 0 841 394 A2

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:214:

GGCACGAGCG	ATGTGCCCAG	CTTTTTC	CAATGCAATACC	ATCGGTCCTT	GATATTGTTT	60
TTCTGGTACG	TCTTTAAGTC	CTAATGTTAA	TTCCACTTCT	TTGTCATCAT	CTGACCATAA	120
TTTTCTGAAA	TCGTGTTCTA	CTAAATCTTC	TTGATATACA	TTTTCTGGTA	ATGGTGGTAA	180
CCATGGACGC	TAACTTCTT	CAATTCTAA	TCGTGTGTA	ATAGATTCTGA	TATGATCGAT	240
AACCGCTTCT	AACTCGGTTT	GGTTTCTTT	CGTTTCTTCA	TCTTCAAGTC	CACTCAAGTC	300
TTTGTGATT	GCTTGAAGTT	GACCATAGTC	ATTAATCATG	TAAATCGTCT	TATCTTCAAC	360
TTCTAATTTA	TCGCCTTCGA	TGTCATATGT	TGCACCACTC	CATGCAGATT	GGAATAATTC	420
ATAAATTTC	TTATTACCAA	CTTTGTAAAT	ACGCACGACC	TGGTAATGTA	ATGCTCGCTG	480
CATCTGGTGT	TTTAAAAAT	TCATTACTGT	CTTGTCTATC	TGGNACGAGT	TNTAATGCCA	540
ATTTAAATTN	AGAGTTAGAC	CAAATTTGGT	CATCAACAAC	ACCCGATGGT	TTNTGTGTCG	600
CAAGTATTAA	ATGAATACCT	AACGAACGTC	CAATACGTGN	CGGNGATACA	AGTTCNTNC	660
ACATAAAATC	AGTTGNTCT	GATTTTAATT	CGGCAAACTC	ATCGGAAATA	ATGAATAAAT	720
GTGGCATTGG	CTCTGNTGCA	ACACCTTCTT	TAAATAACTT	ATGGTATTGA	TTAATATGGT	780
TAACATCATG	CTCTCCGAAT	AAACGTTGAC	GTTTCTCAAT	TCGGCTTTGA	TTGATGTTAA	840
GGCACGCATC	GCTTCATCGC	CATCTAAGTT	TGTAATCCGT	ACCAACTAAA	TGGACTAAAT	900
CTTTAAATAA	GTTCGCCATA	CCCCCACCTT	TATAGTCAAT	CAATAGGAAC	GCAACTTCAT	960
GTGGATGAAA	ATTAATAGCT	AAAGATAAAA	TGTATGATTG	GATAATCTCA	AGATTCCCT	1020
GAACCACTGG	TACCAGCAAC	TAAACCATGT	GGCCCGTGTG	CTTTTTCATG	TAAGTTCAAT	1080
GATAAAATAT	CATCTTTACC	TCTTACACCT	AAAGGTACTG	CCATCGTTTT	GTATGTTTCG	1140
TTTTGTCTCC	ATCGATTAA	CACATCAAGC	TGATCTACTT	CTTTCACGTT	ATACATCTCT	1200
AAAAATGTAA	TACTATCAGG	AATTGCATTT	TTCAAATGTT	CGACGTGTAT	CAAATTCGCC	1260
AAACGTCTCG	CGATGATATT	CTTTATCAAC	GTTATCAATA	TTTTCAGGTG	TAAATTTCAA	1320
TTGAACTAAT	TCTTTTCTT	TCGTAATCAG	TTCCGCTTCA	GTACGAGACT	TGATATCAAT	1380
AATGGTATCT	ACATGCTCTG	GCAAACCTTC	AATCACATCT	TCAACAAAGA	TTAATGAAAT	1440
ACCATATTCT	GATNAATCTT	GGNTNACATA	GTCTAAAATG	ACATGATCAA	TAATTAATGA	1500
CATATCTGTA	ATGACAAACA	CTAATTGCGG	TGTGAAAATA	ATTGCTCAT	TACTTCTGCT	1560
GCGTTCACGC	ACAGCTTGGA	TACGTTCTTT	AATCATGCTA	TAAATTGACG	TTAAAATTTG	1620
GTCACGTGTT	CGGTGATTGT	AAACAAACCC	TCTAATGTTT	TGACCTCTCA	ATGTCATATG	1680
TGGCAACCAA	CGTGCCCAT	TCAATGTTTC	AACTTCATCT	TCACGTGTCA	CAAATAGAAA	1740
CTCTAAATCA	TGATAACTAT	GGAACGTTGA	CAATTGGATT	AGCATTTTCT	CCAATCTTTC	1800
TAAAATGAGA	TGTCGGGCAC	CAATATATGC	AATTGGTCCA	TGATTTAAAT	CATTGATTAA	1860
TGGGTGCTTT	GTCCNACANC	TNTGANAAAT	TCATACAATT	CTTTAGCATC	GTCGAATAGT	1920
TCATCACGAC	GTTGGTTAAA	TTCTCCNTCT	TGGTAATCTA	ATTGAATGA	CTTTTCTACA	1980
TTGCGAATAC	CTAACTTATA	ATGTAAGAAA	TCGTGATGAT	GCGATGTTTT	TTTATATATT	2040
CTTGGTGCTT	TCGTTTCAAC	GATATCTTTA	ATTTCANCAA	CCGTGGGATA	ATGGTAATTC	2100
AAACTAAAAC	GTTGTGCTTT	AATCGCTTTA	TTAATTTCTT	TAGATTTATT	ATCCAAATAA	2160



# EP 0 841 394 A2

... GCTTTGTAAT CTTTCTCTCG TTTTCAACA TCTTTGTTAT ACTTTTCTT TTCAGAGAAA 2220  
TACGTTGTAA TACCAAATAC TATCGTTACT GTACTCATAC CAATCATCAT TAAAATATAA 2280  
5 ATACCAATTG GTCTCACTAA AAAGATGACA ACAGTTAAAG CAATCATTAC TAATGGCGGT 2340  
ATAATGGAAC GCCATATCAC TGTATTCCGG AATCCGGAA TTCCGTATTG GCTGTGGCGG 2400  
TCTTTCAATC TTAATATCAT CGGTCGGGTC NCGGTGNATA ATCCTTGGCG AGCGATGGTA 2460  
CGTATTGTAA TCATCTGCCT GTGCATGTGG CATCTCTTGT GTTAAGCGAA TTAATGACGA 2520  
10 TGCCACTGTA TTCTGACTCA ATACANTTAA ACCATCAG 2558

## (2) INFORMATION FOR SEQ ID NO:215:

### (i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 668 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
20 (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:215:

25 CCGGCTGATT ATTAGCGAAG CGAGATTTTG AAGAGATGCA ACCATTATAT GAATGGTTAG 60  
GTTAACTGC TTCATTAGGT TTTGTTGATA TTGTGGATTA TGAGTATCAA AAAGGGAAAA 120  
ACGTAATATA TATGAACATG ATATTATAAA TACAACATAA GGACGTCTCG GTTTTGATTA 180  
30 TTAAATTGAT AATTGGCTG ATAGTGCAGA AGGGAAATTT TTACCACAAT TAAATTACGG 240  
TATTATTGAT GAAGTGGATT CAATCATTTT AGATGCTGCT CAAACACCAT TAGTTATTTT 300  
GGGTGCACCA AGATTACAAT CAAATCTATT TCATATTGTG AAAGAGTTTG TAGATACATT 360  
GATTGAAGAC GTGCATTTTA AAATGAAGAA GACCAAAAAA GAAATATGGC TGTAAATCA 420  
35 AGGTATTGAA GCGGCACAAT CATACTTTAA TGTTGNAGAT TTATATAGCG AACAAGCGAT 480  
GGTCCTAGTG CGTAATATTA ATTTAGCACT GCGCTCGTGC CGAATGTTCC TTGGTTCATA 540  
TAAATCATAT CGATGTAAAA ATGTTCTTAA ATTTGGGTGC ATGATTGGAG AACAACCAGG 600  
TGGATAATCT TAGTCATTTT CAAGGGTGCA TTCCCATTA ACTTATCATA TCGGCAAATT 660  
40 CCCGCTTT 668

## (2) INFORMATION FOR SEQ ID NO:216:

### (i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 2606 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
50 (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:216:

55

EP 0 841 394 A2

	CCCAAATCCT	TTGAAATAAA	ACTTAATTAA	GTTCAAGAAN	TCGNNGTGAT	TTTCCCCGGA	60
5	AGTTTGTGTTG	AATATTATTA	TGGATTGATT	TGATAGAGCT	TATGGACCCC	ACATTAATAT	120
	TGAAGTCCTG	GGGCCAAGTT	CACNCCCCCG	TGGGTGGAAT	TTCTTATTTG	ATCCCTAAAC	180
	CCAAACTATG	GACGTCCCAA	ATTCTAAATA	TGAAAAAATG	GCTGAGCATC	GTTTGATGAA	240
	TCATGATTTA	TATCCCGAAA	AAATAGATAA	TCGTAAATG	TATCATTAA	TAAACACACC	300
10	AATAAGTTGA	TTTTCCTAAC	TTATTGGTGT	GTGTTTTTCA	TTTAGCATAC	ATAATAGGTT	360
	ACATTAAAAAT	AACATTTTAT	ACCCAAAGTA	CACCAAAAGA	ATATTAGTAC	ACGAATTANA	420
	CAACATTTTT	ATAGAAACCT	ATTGCACTTT	AACGTCAATA	AGTATATTTT	TATATTATCT	480
	CTAATTAATT	GTGCGCGCTT	AATAACAGAA	TATTCTCAAT	ATTTTTATTT	TTTGTGATT	540
15	TGTTGGAATA	TTTAGTGTAT	AAGGCACAAT	CAAAATTTACT	TAAACTATTG	TATTAGGGGA	600
	AGAAAGGATG	GGATGTATAC	ATGACACAAC	AAAACCTCCA	TGGAAATCAA	ATTCAAGACA	660
	TACCTCAAAC	AGGATTTTTC	GGGCATCCTC	GAGGACTAGG	CGTACTCTTC	TTTGTAGAGT	720
	TCTGGGAAAG	GTTTAGTTAT	TATGGCATGC	GTGCCCTACT	CATTTTCTAC	ATGTACTTTG	780
20	CCGTAAACAGA	TAATGGCCTT	GGAATTGATA	AAACAACAGC	TATGTCAATT	ATGTCAGTTT	840
	ATGGTTCATT	AATCTATATG	ACATCCAATA	CCAGGCGGAA	TGGATTGCTG	ACAGAATTAC	900
	AGGCACTAGA	AGCGCTACTT	TATTAGGTGC	AGTCTTTATT	ATTATCCGAC	ATATTTGTTT	960
	AAGCTTACCA	TTTGCAATTA	TCCGCTTATT	CACATCAATG	TTCTNCAATC	TTATTGGCTC	1020
25	AGGTTTAAATG	AAGCCAAACA	TTCCAAATAT	CGGTGGCCGT	TTATATCCTG	AAAATGATAG	1080
	ACGTATGGAT	GCAGGTGTG	TTATTTTCTA	TATGTCAGTT	AATATGGGTG	CATTATTATC	1140
	ACCTATTATT	GTGCAACACT	GTGTTAATGT	TAAAACTTC	CACGGCGGAT	TCTTGATTGC	1200
	AGCAGTTGGT	ATGGCATTAG	GTTTAGTATG	GTATGTACTT	NACANCCGCA	AAAACCTAGG	1260
30	TAGCGGTGGT	ATTGAAACCN	ACTAACCCTA	TGACNACCAG	CTTGNAAAGA	AAAAGTATGG	1320
	TCTTTATTAT	CGGAAGTGGT	GTCTTTAGCA	ANTGTATTAA	TTATCGGTAT	TGGGGCATTAA	1380
	ACTAACTCAA	TATCAANTAA	CTTAGGTAGG	AATACTGTTT	TAGTATTAGG	TATTGCATTAA	1440
	CCANTCATT	ACTTGACTTT	AATTATTAGA	AGTANAGATG	TCANNGATAC	TGNACGTTCT	1500
35	CGTGTTAAAG	CATTTATTCC	ATTATTTATT	CTTGGAAATGG	TGTTCTGGGC	TATCAAGAA	1560
	CAAGGTCTA	ATGTATTAAA	CATATATGGA	ANTGNACATT	CAGATATGAA	ATTAACTTA	1620
	TTTGGTTGGA	AAACANACTT	TGGTGAAGCG	ATCTTCCAAT	CAATTAACCC	ATTATTTATT	1680
	TTATTATTAG	CTCCAATTAT	TTCACTTTTA	TGGCAAAAGC	NTGGAACATA	ACAACCTAGC	1740
40	CTGCCAGTAA	AATTTGCAAT	TGGTACGTTT	TTAGCAGGTG	CGTCATACAT	ACTAATTGGT	1800
	ATTGTAGGTT	ATGCATCAGG	TTCACTCAAAC	TTCTCAGTTA	ACTGGGTTAT	TTTATCGTAC	1860
	ATTATTTGTG	TTATCGGTGA	GCTATGCTTA	TCACCAACGG	GTAATAGTGC	TGCTGTAAAG	1920
	TTAGCACCTA	AAGCATTTAA	TGCCCAAATG	ATGAGTATTT	GGTACTTAAC	TAACGCTTCT	1980
45	GCACAAGCAA	TCAATGGTAC	TTTAGTTAAG	TTAATAGAAC	CACTAGGTCA	AACAAATTAC	2040
	TTTATTTTCC	TAGGTGTTGT	TGCAATTATT	GTTACAACAA	TTTGTATTAG	CATTCTCACC	2100
	TTTAATCATC	AAAGCGATGA	AAGGTATACG	TTAATATTGT	TGGCCTAATT	CAAAAAACAG	2160
	TAAGTCATTT	AAATGGCTTA	CTGTGTTTTT	ATAGGTTTCT	ATTAATTAAA	TTCAAGATAT	2220
50	CAGTATAAAT	AAAAGCTTAA	TATGCTCGTT	ATAGACAGCT	ATAACTATAT	TTTCTCGTCC	2280
	CACTCTATAG	ACATCAATGT	CAGTTATTAC	CATTTTCTCT	ATTTAAAAAC	ATATTTTGG	2340
	TATTAACAAT	TAAAAGTGGG	TATATATATT	TAATGAACCA	ATTTTATAGAA	AGTAGGCTAA	2400
	ACTATGGAAA	ATAAATATAC	ACATGGGTGC	CTCTTTTACC	ATGAACACAG	TGGATTAAAA	2460
	AATATTAATC	AAGGTATTGG	GGAAGTTACA	ACAGCATTA	GTTCAATTG	TAAGCATCTT	2520
55	TCTATTCAAT	TGAGTGAAAA	CGAAGGTGAC	ATTATTAAAT	ATTGCCAAGA	AATCAAAACT	2580

# EP 0 841 394 A2

AAAACTATC CAAAAGATGT AGATAG

2606

## (2) INFORMATION FOR SEQ ID NO:217:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 367 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:217:

CTACGATGGC	ATCTAAATCT	GCGGTGTCGT	CAATAATCAG	TGCTATTGNC	TGGTGGNCCT	60
GCNATGTGAG	CAATACGCTA	CCTGTCCAAA	TAAANATTTT	GNGGNAGATG	CAACAAATTG	120
GGTACCTGGA	CCNACAATCT	TATCAACTTT	AGGTATCGTN	TCTGCNCCAT	AAGTCAATGN	180
AGCAATACTN	TGAGCACCAC	CNNCTNGANA	CACTTGAGTA	NCTTGCGTAA	GATAACATGC	240
AGCTAATACC	TCTTGGGATA	CTCCGTTAGG	GTGAGGTGGT	GTCACANCAN	CAATATTTTC	300
TACACCCGCT	ACTTGTGCTA	AAGTACNCTG	TCATTAGANC	CGTTGATGGA	TAACTTGCTT	360
TGCCACG						367

## (2) INFORMATION FOR SEQ ID NO:218:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1485 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:218:

GGCACGAGCG	ATATTTGTCA	GGAAGTTAAA	GTTTCATGGAG	ATAAGGCTTT	GAAAATGTAT	60
AATCTAACAT	TCGATCATAC	GAAAACAGAT	CATTTAGAAA	TTAGTCATGA	ACAAATTAAA	120
GCAGCATTTG	ACACATTAGA	TGAAAAAACA	AAACAAGCAT	TACAACAAAG	TTATGAAAGA	180
ATTAAAGCAT	ATCAAGAAAG	TATTAAGCAG	ACGAATCAAC	AGTTAGAAGA	ATCAGTGGAG	240
TGTTATGAAA	TATACCATCC	AGTAGAAAGT	GTGGGTATTT	ATGTGCCTGG	TGGCAAAGCA	300
AGTTATCCAT	CAACGGTTCT	AATGACAGCG	ACTTTAGCAC	AAGTAGCGGG	TGTAGAAAAT	360
ATTGTTGTTG	TGACACCACC	TCAACCTAAC	GGAGTATCCC	AAGAGGGTAT	TAGCTGCATG	420
TTATATTACG	CAAGTTAATC	AAGTGTTTCA	AGTTGGTGGT	GCTCAAAGTA	TTGCTGCATT	480
GACTTATGGA	ACAGAAACGA	TACCTAAAGT	TGATAAGATT	GTAGGTCCAG	GTAACCAATT	540
TGGTNGCATA	TGCCCCAAAA	ATATTTATTT	GGACAGGTAG	GTATTGTTCC	ANATTGCAGG	600

EP 0 841 394 A2

	ACCANCAGAA	ATAGCACTTG	ATTATTGACG	ACACCGCCAG	ATGTAGATGC	CANCGNCTNT	660
	NATGTTTTTG	GGCAAGCNAG	AACATNATGA	ATTAGNACGT	CCATNTNTCA	TTGGTGAAGA	720
5	TGCGCAAGTC	CTNAAAGATT	TNGAATCACN	TATTGCTAAN	GCATTGCCTA	ATGTGGACAG	780
	ANACGACATT	GTTTCTANAA	GTATCGCNAA	TCAACACNAC	CTNATCCANN	ATNGTAATTN	840
	NGATNTGGGN	GAAGCATCNC	CANCTCATGA	ATACAATCGC	GCCTGAACAT	GCGTCGATTC	900
	AAACAGTAAA	TCCTCAACCA	GGAATTCTTG	TTGAACGAAT	AATTTCTATC	TCGTTATCAT	960
10	CGTAAGCGTC	ATTATCAACT	TGTGTTTCTT	GCATTTCTTG	TAATTCGGCA	ACAAACACTT	1020
	CTTGATCTCC	TCGATCACGG	CTCTTACGAT	TAATACGTGT	TTTATATTTT	CGAACTTGTC	1080
	TTTCAAGTTT	ATTATTAATT	AAATCAATAC	CTGCGTATAA	ATCATCGTTT	CGCTCTTCAG	1140
	CTCTTAACGT	AACATTTTTT	AATGGAATTG	TTACTTCAAT	TTTAGTAGCT	GAATNTGAAT	1200
15	AAGTTTTAAC	TTTAACATGC	GCCACTGCAT	NTGGTACGTC	ATTAAAATAA	CGTTCCAACT	1260
	TACCAATTTN	CTCCTCAAAT	ATAGNTGCCA	ATAGCATCTT	GTGATAGGGG	AGGGNTATCT	1320
	CCAATGAATT	TNCAAATCTA	TATCATANGG	GAATCTCTCC	CTTAAACCTC	TTTATTGGTA	1380
	ACTCTTTATT	ATATTNNNAC	ATTNNNACGC	TATCGNGCAA	ACGCAAACAC	TTTGGATTCT	1440
20	CTGATATTNT	NNAGCATATT	AATTTACAAC	CCTGCACGAT	GATTG		1485

(2) INFORMATION FOR SEQ ID NO:219:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1514 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:219:

	ATGATGTAAA	CCCAGTTCCN	ANCTTCATCA	TTCAAGTTGAT	GCAATAATGT	CACNNAAACT	60
	CTACCACCAA	GATGCACCAA	TAGGTTGTCC	CTAAATGCAA	TAAGCGCCAC	CCTTTAACAA	120
	TTCAACCTTT	TCAGGGAGGT	NATTTTAAAT	TCACCGATCC	AACAGCTAAT	GATTGTCCTN	180
40	CAAATGCTTC	AATTTAATTC	CAAATACATC	AATATCGCTT	AATTCCTTTT	TACTACGTTT	240
	CCAAAGCCTT	TTCTACAGCG	CCAACGGGTG	CAATACCCAT	AATAGAAGGA	TCTACGCCAT	300
	GACTTCCCAA	AGCCATCCAA	GCACTGCCAA	TGGTTCGATA	TTTAATTCTT	TAGCTTTGTC	360
	TTCTGACATG	ACTAACATCA	TCGCAGCACC	ATCATTGATT	CCTGATGCAT	TACCTNCTGT	420
45	AACTGTCCCG	TCTTTTGTGA	AAGCTGGTCT	TAAGCGACTT	AATTTTTCGA	CTGATACATT	480
	TTACCGGACA	CCTTCATCCT	TAGTGACTAA	GATTGGTTCA	CCTTTACGTT	GAGGAATCGA	540
	TACTGGAACT	ATTTCACTAT	CAAATTCACC	ATTTTGCTGT	GCACGTACTG	CTTTATTGTT	600
	GTGAGATTAC	CAGCAACATG	TATCTTTGTT	CTTCTCTTGA	AATACCATAT	TGCTCCTACT	660
50	AAATTTTCAG	CAGTAATACC	CATATGATAT	TGAGTAAATA	CATCTGTTAA	ACCATCATAT	720
	ACCATGCTAT	CAACCATTGA	GTGATGTCCC	ATTTTAAAAC	CGAAGCGACT	GTTGTTGNCA	780
	AGCATTGGTG	NCTGAGACAT	ATTCTCCATA	CCGCCAGCGA	GCACGATGCT	CATGCCACC	840
	AGTCAGCNAT	AGATTGATAT	GCTAATGAN	TCGACTTTAA	CCCAGAGCCC	ACATACTTTA	900
55	TTCACTGTAA	ATGCAGGTGC	TGTTTCTGGC	AAGCCACCTT	TCATAGCAGC	AATTCGTGCT	960

EP 0 841 394 A2

GGATTTTGTC CTTGTCCTGC TTGTAGTACG TTACCGATGA TAACTTCATC AATCTCACTT 1020  
GGATTCAAAC CCGTCCTCTT AATAATATGT TCTATTAAAG TCGCACCTAA ATCATAGGCT 1080  
5 GGCACGTCTT TAAACGCACC TCCAAAAACG CCAATAGGTG TCCTGTATGC TGCTGCTAAT 1140  
ACGACTCTCG TCATGTCATA TTCATCTCCT TATGATATAT ACAATTTTAT TACTAATACG 1200  
TCTAATGAAC CAATACGCCA CTAACATCCA ATTTATGAAA CCAATTGCTA CGTATGTTAC 1260  
ATTTTCATTC TTGTTATGCA TTTAACAAAT CATATTATAC AATGATGTAA TTTGTTTGGG 1320  
10 CAATCGGAAA CGTAAACGGT TTTACTTTAA CATTTCACAC TCTTAAATAT GCTCTATCAT 1380  
CCTCATCTAT AGTACAAAAT TATGCTTTGA TTTACCACAA ATTATCATTA ATCATTTCGC 1440  
TCGTGCCTCT AATTAAATTC TTTCTATCAT TGGATCGNTA TCACAGGTGG NTTCACCACA 1500  
ATTCTCATCA ATTT 1514

(2) INFORMATION FOR SEQ ID NO:220:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1000 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:220:

GGCACCAGCT CAAGTAAAGT GCTACCATT TTTCTACAGA CATTGCGCA NATCTTAAAC 60  
TAAGACATAN CCCCTAAAAT ACTCNCTANA GNACCGNACA ACCNCAACAC AATTCCTGTC 120  
ATGATCTTAG GCTTTGTGCC ACCTNCCCCA GTTGNCAAAT GCTTTGTCCA TCCTTCTTGT 180  
AATGACTGAA ATCCTTCTTG ATACATNCGA AATGCAACAA ATGGAAATCC TTCATAAACT 240  
35 GTTACGGGCA ATGATTGTGA AGTAGATGCA CTTCTTAAAG CAAATCCTTC AATAAAAATG 300  
ACGGGTGCGA CTTTTATGAC CTCCAGTTGC ATAATAATCT TCTTTATTTG TTAATGTCAC 360  
CGGTCCAAAT GCATGCTGGT TAGTCCGCAC CGTCGTCTAA GGTANAAAAT ACATTCATAC 420  
CACCGACTTG TCATTAAATT AAATATCGCT GAAAACCCCT CGTAGAAACG CTTTGTTATA 480  
40 TGATAAGGCT GTACGCTTAA CAATCCTTTT TCACCTTGTA ATTGATACTG ATTAATCAAC 540  
GTTTCAACAG CATCTTTCTT TAAGAAAGTT ACATCAGCAT CTACAAAGGC AATACGATTC 600  
GTACATGCAT GCGTCACACC TTGATAACAA GCATGTGATT TCCCCATACC ATTGCGCGTC 660  
AGTATCATCA ACAACATCTA CTACAGTCGC ACCATATGAA CGTGCTACAC GTGCTGTTTC 720  
45 ATCTGTCGAT CCGTCATTCA TAACAATGAC ATCTACTGGA ACTTGCTGTT GTATTATCGA 780  
ATGTAGTAAA TGACCTATTC TTTTCTCTTC ATTTCTAGCC GGAATAATAA TTGTTAATGC 840  
TTTATGATTG AAGTTCAGCG TTTTCGTCCT TAGCTGATGT CTACGATTAA ATATCAATGC 900  
ACCACACGCC ATAGACATGG TCACTATTAC TGTTAATATT CGTGATAACC ATTTCATAGA 960  
50 TATCACTTCC TATTCTTCGC TTCTCGCGCG AGCCCGNGCC 1000

(2) INFORMATION FOR SEQ ID NO:221:

(i) SEQUENCE CHARACTERISTICS:

# EP 0 841 394 A2

(A) LENGTH: 1047 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:221:

```

GGCACGAGCT ATATCCAATT GGGGTAAAGC GGGTAAAGAC CAAGAGAAGA ATTTAAGAAG      60
ACATTAGCAG AAATTGAAAG GACACCGGCA CGAGCTAGCG CAACAAGTTT AGCGATTGAA      120
GCATTTGGTG CAAAAGCAGG TCCTGATTTA GCAGATGCTA TTAAAGGCGG TCGCTTTAGT      180
TACCAAGAGT TCCTAAAAAC TATCGAAGAT TCGCAAGGAA CGGTCAATCA GACATTTAAA      240
GATTCTGAAA GTGGCTCCGA AAGATTTAAA GTAGCAATGA ATAACTTAA ATTAGTAGGT      300
GCTGATGTAT GGGCTTCTAT TGAAAGTGCG TTTGCTCCAG TCATGGAAGA ATTAATCAAA      360
AAGCTATCTG TAGCAGTTGA TTGGTTTTC AAGTTAAGTG ATGGATTTAA AAGGTCGATT      420
GTTATATTCG GTGGTATTGC TGCTGCAATT GGCCCTGTAG TTTTGGGAT TAGGTGCATT      480
CATAAGCACA ATTGGCAACG CAGTAACTGT TTAGCTCCA TTATTAGCTA GTATTGTAAA      540
GGCTGGCGGA TTGATTAGTT TTTTATCAAC TAAAGTGCCCT ATTTAGGAA CAGTCTTCAC      600
AGCATTAAGT GGTCCAATTG GTATCGTGT AGGTGTACTG GCTGGTTTAG CAGTCGCATT      660
TACAATAGCT TATAAGAAAT CTGAAACATT CAGAAATTGT GTTAATGGTG CAATTAACAG      720
TGTTAAACAA ACGTTTAGTA ATTTCAATCA ATTTATCCAA CCTTACATTG ATTCCGTTAA      780
AAACGTCTTT AAACAAGCGG TTTCAGCAAT CGGTGATTTC GCTAAAGATA TTTGGAGTCA      840
AATTAATGGA TTCTTTAATG AAAACGGAAT CTCTATTGCA CAAGCGCTTC AAAATATATG      900
CAATTTTATC AAAGCTATAT TTGAATTTAT CTTAAATTT TGTAATTAAA CCAATCATGT      960
GTGCGATTTG GCAAGTGATG CAATTTATTT GGCCGCGCGC TAAAGCCTTG ATCGTCAAGT     1020
ACTTGGGAGA ATATAAAAGA GTAATAC                                     1047

```

(2) INFORMATION FOR SEQ ID NO:222:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1526 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:222:

```

GGCACGAGCG GTACGAGTTC ACTGTTAATC TGCACTTCAA TGTGATTATG ATCTGAGTTA      60
TCCTGCATTT CGTAACCGAT AGCGATGTTT GTTACTTTTT CAAAATTAAT ATCACTATTG      120
TTCGCTAAGC TTGAACGTAC AGACTCTAAT TTCGGCAACG TTTTATTATC CTCACTATTT      180
AACACAACGT CGGTGCGCAA TAACGAACGA CGATAATCAA AGACGCCTTT TTCACCCCAA      240

```

EP 0 841 394 A2

	GTGACTTGAA	TTTGATTAGA	GCCTTCTTTA	TTAAACGTTG	GATAACCGTT	AAGGAAACGT	300
	TGATATGTTA	ACTCGCCTGA	CTGATGGCAT	TCGTACTAAA	CAATCTGAAG	TCTTCGTTTA	360
5	AGAAACCACC	ATGACCATTA	ATAAAATCAA	AGGTTCTTGG	AATCGTTTCT	TCCATTGTGC	420
	TGGAACCTGC	TTCATCTTCG	GACAGGTTTT	TATAATGATA	TTTCNCATTT	TTATCGTTAT	480
	AGTGTGNGAC	ACCTGTATTA	TTGNTGTAGG	TTGTAACACC	ACTCTNTGAN	CTACGAACGA	540
	TGGNTGANTC	GACAAATAGT	ATAGCATNCA	TTTTCTCAAC	ACTAATCGTG	TTAAATACCA	600
10	TGCGATATGT	TTTTANCTTN	NCAGGTTTAC	CTTGGGTGCA	AAAACATGCG	TCGNACGATC	660
	AATTGTATCT	NTGTGTGTGA	TGATATCTTG	TGTATGGGTT	GCATATCTTT	TCACACTGCT	720
	GCTTAATGCA	TCTTAAAAAA	TGNATCATTT	CGCGGTGGTA	GNTGNTTAAT	TTTACGGTAA	780
	TCGGGGCGGT	CTTACTTTAT	AGCATAAAGC	ACGATATTAT	CATCAGCATC	ATGATCTATG	840
15	ACTAAACGAT	TGAAATTGAA	ATGATTTGGT	ACTTTCGCAT	TCAATGTTCA	NGAACTTGGC	900
	CAAGATATGT	TGATAACGGT	AAATCAAACG	TAAAAATCGAA	TAATGTAAAA	TCACTGCTTA	960
	AATCCGGAAT	CATCAAGTTA	TGATCACGAC	GAACATGTTT	CACGGACNTC	ACTTCTTTAT	1020
	TTTTCAAAGG	TTTCGTCAGT	TTATTFCACAT	TAGATACCGT	CGCAATCGTA	CCTTCTGGAT	1080
20	GATCATTTTT	CGAATGAATA	ATCTGAAATG	GCGTAATAGT	TGTATCCATT	TTGGCTGTCA	1140
	TAGGTGTCGT	TAAAGGTTTC	GTTTCACTCT	TCTTACTATC	TGTATTGTCG	ACATTTGCAA	1200
	TATCAGGAGA	AAAGTTCCAT	ACCATATATG	TCAATACGAC	ACTCATCAAG	ACGAGTAGCG	1260
	CTAAAATGAC	AGATTTAATA	TGTTCCTTAT	TATTCATCCC	AATCACCGTC	TTCAATGACT	1320
25	TCACATGGAA	GTGTGATAAA	GATAGATGTA	CCTTGACCTT	CTACACTGTT	TGCCCAAATA	1380
	CGACCATTGT	GCGCTTCCAC	AATCTCTTTC	GAAATGGCTA	GTCCTAATCC	AGTACCACCC	1440
	ATTTTACGCT	CGTGCCGTNC	CTTATCTACA	CGATAGAATC	GGGNCGGAAT	ATCTTTATCG	1500
	ACTTTATTGG	ATAGGAATTC	CCGATG				1526

(2) INFORMATION FOR SEQ ID NO:223:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1421 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:223:

45	GGCACGAGCG	AGCAATATCA	AATAATAAAT	CATCATCAAT	AGTTCCTCCT	TGGAAACGTT	60
	CAGGGATGAT	GTTGAATAAT	AATGACGTAT	CTAATATATG	GTCATATAAA	GAGAAATCAC	120
	CAACTGGGAT	GCTATCTAAG	TGATAGTACT	TTTGTAATAA	TAAATTTTCT	TTATGTAGAT	180
	CAGGNAATGN	ATGCATCTAN	CNACTCNTCN	TTAGANNTCN	TCNTTGCCCA	ANAACTTTTCG	240
50	ATGGCTTTTC	TCCATTCCCT	TTCTCNACCT	AATCTTGGGA	ANCCTAAGTT	TGATGTTTTA	300
	ATTGTTGACA	TAATATTGCC	TCCTTGTTAG	CAGGTAATAG	ATTTNGAGTA	TGCTGCAAGT	360
	TCTAATGAAT	CTTCGACATT	TNGANACGGT	GTGATAATGT	ATAAACCATN	AAAATATTCA	420
	TGAACAGTAT	CGATTAAATC	CTTGAAAGC	TTAAGACTTA	GTTCTCGCGT	TTTGGCTTTA	480
55	TCATCTCTAA	CTGCTTCAAA	TTGGTGAAAA	ATTTCANCTG	ACATCTTGGA	TTCCNGGCAC	540

EP 0 841 394 A2

	TNCATTATGC	AAAAAGAGTG	CGTCNTTGTA	ACTTGCGATA	GGCATAATGC	CTATGAAAAA	600
	TGGNTTGNTC	AAGTGCTTAG	TGGCATGGTA	AATTTCAATG	ATTTTCTCTT	TGCTGTACAC	660
5	GGGTTGTGTT	ATNAAATAAG	ACATTCCGCT	TNCTATCTTT	NTNCTCTAAT	CTTNNGACGG	720
	CACCATCTAA	TTTACGAACA	TCAGGGTTAA	AGGCGCCAGC	GATGTTGAAG	TGTGTACGTT	780
	TCTTCAGCGC	ATCACCCCCA	GNGTTAATAC	CCTGATTAAA	TCTTAGAGCG	AGTTCANCNA	840
	ATCCTTTAGG	AATTAACATC	ATAGACATTG	GTTGCACCTG	GTAAGTGACC	AACTTTGGAA	900
10	GGATCACCAG	TTATGGCTAA	TATTTTCGTTA	ACGCCAATGA	GCGATAATCC	AAGTAAATGG	960
	GACTGCAAGC	CGATTAAGTT	TCGGTCTCGA	CATGTAATAT	GTACGAGTGG	NTCAATATTG	1020
	TAATATTGCT	TAATTAAGCT	AGCAGCAGCA	ATATTGCTAA	TTCTGACAGT	TGCCAATGAA	1080
	TTATCTGCGA	GTGTTACCGC	ATCTACATTA	GCTTTTATCAA	GTTTAGCGAT	ATTTTCAAAA	1140
15	AATCTATCCG	TGTCTAAATG	TTTCGGTGTA	TCCAATTCGA	TAATAACGGT	TGGACGTTCT	1200
	TGAACCTTAG	ATGTTAATGA	TTGTCTAACT	TTATTTTGAG	ATGGATTGAA	AAGTGCTTTC	1260
	GTTGGTATCG	GAATCACTTT	TNTGNCANTA	ACAGGTTTAA	GTGTCNGAAT	AGATTCTTTA	1320
	ATAAATTTGA	TGTGCTCTGG	CGTTGTACCA	CAGCAACCAC	CAATTAAACG	AACACCTTCG	1380
20	CGAATTAGTG	CCATTNTGGN	GNAACTTGAC	CGAAATATTG	T		1421

(2) INFORMATION FOR SEQ ID NO:224:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1886 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:224:

35	TTATTTAAAA	GTATCCAATA	GCCAAACAATC	TCTTTCATCT	TTGAGTAGTA	TCCCACGTAT	60
	TTCGGGATTG	CATCCGAGTT	GATTTTATN	ATTTTCTTCA	ATGCTATCAA	GCGCAGCTCG	120
	TTTACGTCTT	ACTTTAGGTT	TTATTTGCTC	AATTGCCTTG	ATTGTTTGAT	TTCTAACATC	180
40	AGTAACAGCA	GCATCTTGAT	TTGTATTTTC	TATTTCTTGT	TGCGCTTGTT	TGAGTGTGTC	240
	GCTAATTAAT	TGATTTCGCT	CATCTAATTC	ATCAACTGTT	GCATGTGGTG	TATCTTTTAT	300
	TGATTCTACT	TGATTTTCTG	CAGTTGCTTT	TATTGCTTGT	TGTGCTTCAG	GCTTAATTAC	360
	AATATGAGGT	TGCACGAGCA	CCTTTTAGTG	TAGCAATGCC	ATTNGTTTCA	ACACGTTTCA	420
45	CATCATTANT	CGTGA CTGCT	TGATTGAGGT	TTTGTAATGC	AAGTTTTC	TTATTCGCTA	480
	ATTGATTTAA	AGCAACTTGT	TTTCTTTCAT	CAGTCNCATG	TTCAGCTTGC	TCTATTTCTT	540
	GCTTTTTAGC	CTCATATTGT	TGCTTTACTG	CATCTCNAGC	AGCTGCTCTA	AAAATATGTT	600
	CAGGCGCTAC	TAAAGCAATG	CTATCAAGCG	CTTGACTTGT	TGTATCATCA	ACTTGTTGAT	660
50	TTGTCTATT	ATTTCGTAATA	TCTGTCAATG	CTTGATTTAC	AAATTCNTAT	GNTNTTATCT	720
	AGTGCTACGT	TGTCNNCTCNG	CCTGCTGTTG	CNTCTTTATC	CCTGANTAAAT	CGTAGCACAG	780
	TAAATCATTC	GCNTGTTGGA	TTGAGTTTTT	CCACGTTGCC	AGCTGGNGGT	AACTTTTTGT	840
	TTCAGGGTGA	ATAATNTTAA	TCGCTGATAC	ACCATTTGTT	CCTGCTTGAT	TCACCTGACN	900
55	ATTTCGTTTGA	GCNTGGTCAA	TANCTCCAAG	TGCTTTTCT	TTTTCTTTAG	CTAATGCTTG	960



EP 0 841 394 A2

5 TGAAGCAACT TCTTCTCAT NATCTGTTGA ATCAAGACTA TTATCAATTT GCTGTTGCTT 1020  
 TTCTTTAACA GCTTTTTTCAA NATCTGCAAT TGCCTTTGGT TTAATTACTA CTTCAGCTNC 1080  
 AACATTATCT ATAGCATTTA CCGCTTGATT TGTAAGTTTG TCTACCTGAT CATTTGTTTG 1140  
 GTTTTGATTA ATTTGATTAA TTGCTTGATC TTTAAGTTGA GTGATTGAT TAACAGCAGC 1200  
 CTGCTTTTCT TCGTCAGTTG CATTAGGTGT TTGTTTAACC GCTTCAATAC GCTTCGCCAC 1260  
 TTCAGCAGTG ATTTTATCTC GCGCTGCTTG TTTTTTTACT ACATCAACTT GAACAGCATC 1320  
 10 GATATTATTC TCTGCTACTG TCCGAGCTTG GTCTACTTCT GCATTGTGT TAGCTTGTTT 1380  
 AATACTTCNA ATAGCTNGTT GTCTGTCTTG ATTTAAAGTA TAGATCGCAG CATTTNTCTC 1440  
 AGCATTCTGT GCATCTGGTG TAGCATTGAT TTCAGCTAAT TTAGCATTAT AATGCTGATT 1500  
 GATTGTGCT AATGCTGCAG GTTTATTAAAC AATATTGGC TGAATCGCAT TAATTGCTTT 1560  
 15 TGTACCTAAT TGTTCGCTT GATCTACTTC CGCATTGTA TCAGCTTGAT TTATATTATT 1620  
 AATTGCCGTT GCTAACTCTT GATCCACTTG ATTTAAAGCC ACTTGCTTTT CTTCAGTTGT 1680  
 TGCATTGTG TTTTGATTAA TTTCTTGCTT TTTAGCAGTT GCTAAATCAT TTAATACACC 1740  
 TGTAGCAGGT TGTTCTTCG TTACATGCGG TTGACTGCGC CGATTGATG ACTGCATCGN 1800  
 20 CTCTAATACT ATTGACCACC GAGTAGTAGA CGTCACACAA TATCAGTAAT GCTCTATTTT 1860  
 AGTGATTGAA ACGAATATCG GTCTTG 1886

(2) INFORMATION FOR SEQ ID NO:225:

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1381 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 30 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:225:

GGCAGGAGAC AATATATCTA CCGTGAATAA CTGTGCAAAT AGTGAAGTTG TAGCCGCCAT 60  
 ACGCATTTCA TTTTCATCAG TTCTGCCATA AATCAATGCA TAGTCGGCAA TTTGAGCCAC 120  
 40 GTGCCGATTA TTCGGCACGA GCTGTACTAG ATATAGTTAT GATGGAATA CTGTAATGTG 180  
 TGGCCACCTG TGCAATTGAC TGCAATTCAC TATGACTACC TTGATTCGTC ACAAAAATCA 240  
 TGCAATCTCT ATCATCATGC TCGCAAAATG TTGACACAAG TAAATGCGTT TCATGTAATA 300  
 ACCTGACATT TAAGCCAATA CGAGATAACT TTTGAAAAAG ATCACCATAA GTCAAACTCG 360  
 45 ATGCGCCAAA TCCAAATAAA AATATTGTCC TGGCATTFTT CAACACATCA CAAATTGCAT 420  
 CAATTGCGC ATCCATAATA TTAGTAGCTA CAAATCGCAT CCGTATTCCG TTGCTCTAGC 480  
 AATCATTTTA TTTTTCAAAG TTTCTACAGA TTCATTTTCA ATCAATTCTA AATGTGGATT 540  
 GGTGCAATA TCTTCGGGTA AGTATCGAGA TATCGCAATC TTTAGCTCTT GAAAACCTTG 600  
 50 ATGTGTCATT TTCCGACTAA ATCTAACAAT TGATGCTGTA CTAACATTCTG TAACATCTGC 660  
 CAAATCATTC ACCGTCATAT CAATGATTTT ATGTGGATTC TTTAAATGT AATCAGCGAT 720  
 TATCTTTTCT GTCTTCGTAA AATCACTCAA CTGTTTATCA ATGCGATATT AAAATATTTG 780  
 TCATCATTA TACCCCAACA AATCTGTCTT GTCGCATCGC CTTTGTGCGT CCAAATAAAT 840  
 55 ATGTACAAAC GAATCCACCA GCATATGCAG CAAGTAATCC TGCAATATAA CCTAAATACA 900

# EP 0 841 394 A2

	TATTATCTGA	GATTAATGGT	AAGACTGACA	CACCACTTGG	GCCTATTGCT	TTGGCACCAA	960
	TATGTCCAAT	TCCACCTATT	ACAGCGCCAC	CAATACCACC	ACCAATACAA	GCAGTTAAGA	1020
5	AAGGTCGAAC	TAATGGGCAA	AGTCACACCA	TAGATTAATG	GTTCTCCGAT	ACCTAGGAAA	1080
	CCAAGTGGCA	ATGCACCTTT	TAAAGTATTA	CGTAATGTTG	TGTTGCGTTT	ACATCTTACC	1140
	CAAAGTGCTA	ATGCGGCACC	TACTTGTTCCA	GCACCCAGCC	ATCGCTGCAA	TTGGCAATAA	1200
	GTAAGTAGCA	CCAGATTGGT	TAATCAATTT	CTATATGAAT	TGGCGTNAAA	AATATGATGA	1260
10	AGGCCCTAAC	AATAACTTAA	CGGTAGGAAG	CTTGGTCCAA	ATGATAAATC	CACTTAAATT	1320
	ACGNCACCCA	ATACTTAATA	ATCCCCGTTA	CTACTGAAAC	TTAANTTGTC	TTTGAAACAA	1380
	A						1381

15 (2) INFORMATION FOR SEQ ID NO:226:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 724 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:226:

30	GGCACGAGCG	TTTGTAGTTGA	TGTGTATCAC	TAATATCTTT	GAAAATTTTA	ATCAGGTACT	60
	ACGACAATAT	GATGTCTGTT	TTGTGTCTGA	AAGTTTACAA	GTTTTTAAAA	TAAAAATGGT	120
	ATAAAGTGTG	ATTTGTATAA	AAAAGAGTCT	CGACGGATAA	GAATTGATTA	ATAACAGTTA	180
	GCATTTTATT	AATTACCTTA	ACAATGATTC	AAGTTTAGTT	AAATGAGGTT	TAATTTGAAA	240
35	GGGATAGCG	CCTCAATATA	ATGTAGGTAG	ATTGTTTCATA	TTACGTAATT	GAAAAATCAA	300
	ATTTAAATAN	ATAGGGTGGG	GCTNNAAATT	ATGAAAATTA	AAGCGATAGC	AAAANCAAGT	360
	TTAGCATTGG	GAATGTTAGC	AACAGGTGTA	ATTACATCGA	ATGTACAATC	AGTACAAGCG	420
	ANAACAGAAG	TTANACAACA	AAGTGAATCA	GAGTTGANAC	ACTATTATAA	TAAACCGGTT	480
40	TTAGAGCGTA	AAAATGTTAC	TGGATATANA	TATACTGAAA	AAGGTAAAGA	TTATATAGAT	540
	GTCATAGTAG	ACAATCAATA	TTCTCAAATT	TCTTTAGTTC	GGATCTGAAT	AAAGACANAT	600
	TTNAAGATGG	AGACAACCTG	NATATAGATG	TGTTTATCCT	TAGAGAAGGT	GACAGTAGAC	660
	AAGCAACNAA	TTACTCAATT	GGTGGCGTAA	CACAAACAAA	CAGTCAACCT	TTTATTGACT	720
45	ATAG						724

(2) INFORMATION FOR SEQ ID NO:227:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 618 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

EP 0 841 394 A2

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:227:

5

10

15

20

```
CCTTCACCAA TACGTGTANT ATCNTTATTT GGGCGNTTGT CCACTTACCC ATGTAATCGT    60
GGCATCTGGA ATGGCACTAC CATTAGATAA CTAAAGTAA TCTGCTGCAC TAGAACCATT    120
ACCTGCTGGG AAGTCTTGGC CTTGTACAGT GTAATGCGAA TGTGCAACGA TTTCTGGAAT    180
GACATGCTTT ACAGGTACAG TCACTGTTGA TGTTGTTCCA TCTTGATGAG TAACAGTTAC    240
TGTGACATTG GCATTATTTT CAGTACGACT GACATTACTA ACTGTTTTAC TAGTGATTTC    300
ATTTGCGCTT GCTGTTGCAT AGTTTCTATT TGGTGCATTA CTTGTAAACG TTAACGAATT    360
AATAATTGTC GCTTGATCAG CTTGTGATAC TGTCGCATTA TTCGAAATAT TGGCAATTCT    420
AACAGGATTA GCAGNCCGTT GATGAAGTAC CAACTCGATA TTATCACGC AAAGGTTTCA    480
CTGTTACATT GAATGAAGTT GTAGCTGTAT TACCACNGT ATCAGTTGCT AATAAATTGA    540
TTGTCTTACT AGTTGCTGAT GTCACATTTG GTGCCGTTGC AGAAACATGT TGATGGTTAT    600
TATCAACAGT CTCGTGCC                                618
```

(2) INFORMATION FOR SEQ ID NO:228:

(i) SEQUENCE CHARACTERISTICS:

25

(A) LENGTH: 1005 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:228:

35

40

45

50

55

```
TTGTGAGACA CACGCTTAGT GCACATGACA ATGTGAATGT TGTGACAA ATTATTAAAA    60
GAAGACTTAC CACTGTAAAG TTGTATCGAG CCGAATGATA CATTTGATT TCCACNAAAC    120
TAGGGAGTGT AAGAAGTGAT GGAATAGT AGGCCCGAGC GTAATGAAGC GACGATGCAT    180
CTTGATGAAA TGACTGTGGA AGAGGCTTTA ATTACGATGA ATAAAGAAGA TCAGCAAGTC    240
CCGTTAGCAG TTCGAAAGGC AATACCACAA TTGACAAAAG TAATAAAAAA AACAATTGCA    300
CAGTATAAAA AGGGTGGACG ATTGATCTAT ATCGGTGCAG GTACAAGTGG AAGATTGGGT    360
GTCTTAGATG CAGCGGAGTG TGTACCTACA TTCAATACTG ACCCTCATGA AATTATAGGT    420
ATTATTGCTG GTGGACAACA TGCTATGACG ATGGCTGTAG AAGGTGCGGA AGATCACAAA    480
AAATTAGCGG AAGAAGATTT GAAAAATATA GATTTAACAT CAAAAGATGT CGTTATAGGA    540
ATTGCCGCGA GTGGCAAAAC GCCATATGTT ATAGCGGTT TAACATTTC TAACACAATC    600
GGTGCTACAA CAGTATCTAT TTCATGCAAT GAACATGCAG TTATAAGTGA AATTGCGCAG    660
TATCCAGTAG AAGTTAAAGT TGGTCCAGAA GTATTAACTG GTTCANCACG TTTAAAGTCT    720
GGTACAGCAC AAAAGTTAAT TTTAAATATG ATTTCAACCA TCACAATGGT TGGTGTGCGA    780
AAAGTTTACG ATAACCTCAT GATTGATGTT AAAGCAACCA ATCAAAACT GATCGACCGT    840
TCAGTGCCTA TTATTCAAGA AATATGTGCT ATCACATATG ATGAAGCAAT GGCGTTATAT    900
CAGGTATCTG AGCATGATGT TGAAAGTTGC GACAGTTATG GGTATGTGTG GCATTTCTAA    960
```

# EP 0 841 394 A2

GGGAAGAAGC AACAAAGACGG TTATTAAAC AATGGTGACC TTGTT

1005

## (2) INFORMATION FOR SEQ ID NO:229:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1040 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:229:

GGCACGAGCG	ATGACAGGAA	TATGATATTG	TTTTGCAGCT	AACGCAACGC	CTACGGGTGT	60
TTTACCAAAG	ATGGTCTGAT	AATCCATGCG	TCCTTCTCCA	GTAATAACGA	GGTCTGCATC	120
TTTAATCTT	TGATGAAAAT	CTGTAATGTC	AAAGACGACA	TCAATACCTT	TTGTTAAAGT	180
TGTCTCACA	AACGCTAATA	ATGCTGCGCC	CATACCGCCA	GCTGCACCAG	AACCTGGTAT	240
TTGATTATAC	GGACTTTCCT	GTGCACATTT	TTATCTTATC	ATGACAATGN	GACATTGCCA	300
AATCCAAATT	AGGTATGCAT	CTTTCATCA	GCGCATTTTT	GAGGACCATA	AATATAGGTA	360
GCACCATTTT	CACCCAATAA	AGGATTTGAA	ACATCACAGG	CCACTTTAAA	GGTCACCTCT	420
TTTANTCGCG	AATCGGCCAG	ATTGGTTATA	TCGATTTGTG	CAATGTGAGC	AAGATTAGCA	480
CCATTCATTT	GTAATAAGTC	CCCGTTTACA	TCAGTAAACT	TTACGCCTAG	TGCACTTAGC	540
ATACCTGTAC	CACCATCAAT	TGTTGCACTG	CCACCAATCC	CTAAAATAAT	GGTCTTAGCA	600
CCATGATTTA	ATGCATCTTT	AATTAGTTCA	CCGGTACCAT	ATGATGATGT	GTATAAAGGA	660
TTACGTTTCT	CTTTTCTTAA	TAAATCCAAA	CCTGACGCTG	CCGCCATTTT	AATAATTGCA	720
ATTTGTTGTT	CGTCTGCGCG	TGCATAACAT	GCTTCAATAG	GTGCGATTAA	AGGGTCATTA	780
ACGATGACTG	TATACTTAGT	GGCANCTGNT	GCATGAATTA	ANGCATCTGN	GGTACCTTCA	840
CCACCATCAG	GCATCGGAAT	GATATCATAA	TGAAGGGTAT	TCCCATAAAC	ATTAGTAAAT	900
GCCTGTTTTA	TAATATTGNC	AAC TTGGTGT	GNGGNCATAC	TTTCCTTAAA	TGAGTCAGGA	960
GCGATGACAA	TTTGTGTTGGA	ATGGATGGNA	GNCACCNCTC	AATATGNNGG	CGNTCTTTTG	1020
GTTTTANATT	NTAGTGACTN					1040

## (2) INFORMATION FOR SEQ ID NO:230:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1291 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:230:

EP 0 841 394 A2

	TGTATAGGCC	TATTTGTATA	NAAAGTAATA	TGATGTTTAG	TGGCTATGTA	AGTGAAGCAA	60
5	TGTATAAATA	ATGATGAGTG	GTTTGTACT	AATGATAATG	GCTATGTAAA	AGAGCAGTAT	120
	TTATATTTNG	CGGGACGTCA	ACAGGATATG	TTAATTATTG	GTGGGCGAAA	TATATATCCA	180
	GCACATGTTG	TNCGCCTTTT	AACGCAATCT	TCGAGCATTG	ATGAAGCAAT	TATCATCGGT	240
	ATTCCACATG	AGCGTTTTGG	TCANATAGGC	GTATTGCTTT	ATTCTGGTGA	TGTGACACTT	300
10	ACACATAAAA	ATNGNAAACA	ATTTTNNAAA	AAGAAAAGTG	AACGCCATGA	AATTCCATTTC	360
	GATGATTCAT	CATGTAGAAA	AGATGTATTA	CNCTGCAAGT	GGTAAAATTG	CTAGAGAAAA	420
	AATGATGTCG	ATGTATTTGA	GAGGTGAATT	ATAATATGAA	TCAAGCAGTC	ATAGTTGCAG	480
	CTAAACCGAA	CTGCATTTGG	GNAATATGGT	GGCACTTTAA	AACATTTAGA	GCCAGAACAA	540
15	TTGCTTAAAC	CTTTATTCCA	ACATTTTAAA	GAGAAGTATC	CAGAGGTAAT	ATCTAAAATA	600
	GATGATGTAG	TTTtaggtaa	TGTTGTTGGG	ANTGGTGGCA	ATATTGCAAG	AAAAGCATTG	660
	CTTGAAGCGG	GGCTTAAAGA	TTCAATACCT	GGCGTCACAA	TCGATCGGCA	ATGTGGGTCT	720
	GGACTTNGAA	AGTGTTCaat	ATGCATGTCG	CATGATCCAA	GCCGGAGCTG	GCAAGGTATA	780
20	TNNTGCAGGT	GGTGTtTgan	AGTACAAGTC	GAGCACCTTG	GAAAATCAAA	CGACCGCANT	840
	CTGTGTACGA	AACAGCATTa	CCNGAGTTTT	TATGAGCGTG	CATCAITTTG	ACCTGAAATG	900
	AGCGACCCAT	CACATGATTC	AAGGTGCTGA	AAATGTGGCC	AAGATGTATG	ATGTTTCAAG	960
	AGAATTACAA	GATGAATTTG	CTTATCGAAG	TCATCAACTG	ACAGCGGAAA	ATGTAAAGAA	1020
25	TGAAAATATT	TCTCAGGAAA	TATTACCTAT	AACCGTTAAA	GGAGAAATAT	TCAATACTGA	1080
	TGAAAGTCTA	AAATCACATA	TTCCGAAAGA	TAACTTTGGC	CGATTTAAGC	CCGTAATCAA	1140
	AGGTGGGACC	GTTACCGCTG	CGAATAGTTG	TATGAAAAAT	GATGGTGCAG	TTTTATTGCT	1200
	TATTATGGAA	AAAGATATGG	CATACGAATT	AGGTTCGAG	CATGGTTTAT	TATTTAAAGA	1260
30	ATGGTGTTAC	GGTAGGTGTT	GATTCTAATT	T			1291

(2) INFORMATION FOR SEQ ID NO:231:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 557 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:231:

	CATTCACTCTA	ACCTNNTTCAT	CAGTTGAGTG	ATTTTAATCT	ATCCTTCAAA	AGTTNGNGGA	60
	TCAGGTCCAA	TTCTCTTATC	TAGATTTAAA	CCAGGCACGA	GCGTCATTTG	TTCATCTGAT	120
	AATTCGAAAT	CAAATATTTG	GAAGTTTTCa	GAGATTCTGT	TTGGTGTtAC	CGATTTAGGG	180
50	ATTATAACCA	CACCATGCTG	CACATTCCAT	CTTAAAACAA	CTTGGGCAGG	TGACTTTCCCT	240
	AATTCTTGAG	CAATGTCTTT	AATGTCTCA	TCATTTAAAA	TTGTGTCATT	CATCAATGGT	300
	GACCAAGATT	CCATCACGAT	ATGTTGTGCT	GCCAAATATA	ATTTCAATTT	ATGTTGCGTT	360
	AAATATGGAT	GATATTCAAC	TTGATTAATT	ACAGGTTTAA	TTGACACTTG	TGCCAACAAA	420
55	GCTTCCAAAT	GTTCAAGGTC	AAAAATTGCT	GACACCTATA	TTTTTAGCTT	TATTATTTTT	480

# EP 0 841 394 A2

ATATAAATCT TCCATACCTT TCCATGTTAT CAACCATTTAC GGCTTCGTTC GTGCCTGGCC 540  
AATGTTACTA RRDCNTS 557

## (2) INFORMATION FOR SEQ ID NO:232:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 904 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:232:

GGCACAAATTC	GGCACGAGCG	TTGGTATGTA	GAATATGATG	GTGAATGGTA	TGTGTATAAC	60
GATGGGAGGC	TTGAATAAAT	GAACGGAAGA	CTGACAAAGA	CACTTTTCAT	TTTCGTGTTT	120
ATTCTTGTC	ACATCGGGTT	AGTATCGATT	TATGTTAATA	AAGTCAATCG	CTCACACATT	180
AATGAAGTTG	AGAGTAACAA	TGAAGTTAAT	TTTCAGCAAG	AAGAAATTAA	NGNACCGGCT	240
AGTATATTGA	ATAAATCAGG	TAAAGGTATA	AAATGAGAGC	AAATTACAGG	GCGATCAAAA	300
GACTTTAGTT	CTAAAGCTAA	GGGCGATTCT	GATTTGACCA	CATCAGATGG	TGGAAAATTA	360
TTGAATGCGA	ACATTAGTCA	ATCGGTAAAG	GTCAGTGACA	ATAACTTAAA	AGATTGAAA	420
GATTATGTTA	ACAAACGTGT	GTTCAAAGGT	TCAGAATATC	AATTAAGTGA	AATTAATTCT	480
GGTTCTGTAA	AATACGAACA	AACGTATGAT	AATTTCCCGA	TTTTGAACAA	TAGTAAAGCG	540
ATGTTGAACT	TTAATATAGA	AGATAACAAA	GCGGCTAGTT	ATAACAATC	AATGATGGAT	600
GACATTAAGC	CCACAGATGG	TGCAGATAAG	AAGCATCAAG	TTATTGGAGT	GAGAAAAGCA	660
ATCGAGGCAT	TATAGTATAA	TCGTTACTTG	AAAAAAGGTG	ATGAAGTCAT	TAATGCTAGA	720
CTCGATTACT	ACTCAGTCGT	GAACGAAACG	AATGTTCAAT	TGTTACAACC	AAACTGGGAA	780
ATTAAAGTGA	AGCATGACGG	TAAGGACAAA	ACGAATACTT	ACTATGTCGA	AGCGACAAAT	840
AATAACCCTA	AAATTATTAA	TCATTAATAT	GAATCGTAAT	AAGCTAGTAT	TGCAAGCTCG	900
TGCC						904

## (2) INFORMATION FOR SEQ ID NO:233:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 699 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:233:

# EP 0 841 394 A2

	TTGTACCAQC	ACAGTTGTCA	CGGAGACCAT	AAAGATGCCG	AGGAGTAAAA	GTAAAAAGAA	60
	CCAAATAGAA	TTGATACCAG	CGGCTGATAT	GTAAACTTG	TTAGGGTTTG	AATTACAACC	120
5	AACTAATGAT	GGATTGATTA	TTCATCCGTC	AAGAATTTAA	AACAAATGCA	ACAGTTGATA	180
	GTTTAACTGA	TCATCGAATA	GGAATGATGC	TTGCAGTTGC	TTCTCTACTT	TCAAGCGAGC	240
	CTGTCAAAAT	CAAACAATTT	GATGCTGTAA	ATGTATCATT	TCCAGGATTT	TTACCAAAAC	300
	TAAAGCTTTT	AGAAAATGAG	GGATAATATA	AAATGGAAGA	TATCTATAAA	TTAATAGACG	360
10	ATATCAATCT	ACAAAAACTA	GAAAATTTAG	ACTCTCGTGT	TAATGAAGCA	ATAACTACTG	420
	ACAACGATGA	CGCATTATTT	ATTCTAGGAG	AGACACTTTA	CAATTTTGGA	TTAATGCCAC	480
	AAGGTTTGGA	AGTATTCCGC	TCGTGCCATA	TCNCAAATAT	CCAGNCGANA	GTGANGTGCT	540
	GATTTATTTT	ATTGAAGGTT	TAATGTCTGA	NAATCAACCT	GCCGAAGCGT	TAGAANANTT	600
15	AAGTTATGTT	GATCCATCAC	CTGCNNAGTT	GATGTTNAAG	AAATAGTTTT	GCGANNTGAT	660
	TATCTAAAAC	AGTATTCAGA	ANTTATTTGG	TAGAAAATA			699

## (2) INFORMATION FOR SEQ ID NO:234:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 654 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:234:

	CCAAACAACC	TATCCTTCAA	ACAATTATCC	AATATTTGTT	GAACATGGTG	CAATTGACCA	60
	TATTAGCACG	TATATTGATC	AGTTTGATCA	AAGTTTTATA	TTAATTGACG	AGCATGTAAA	120
35	TCAATATTNT	GCTGAAAANT	TTGATTGAAA	TTTTAATCAA	ATTGAAAATG	TCCATAAAGT	180
	TCATTAATCC	NAACCTGGTG	AAAAAGACGG	AAAACATTTG	ACCAATATCA	AAGAAACATT	240
	AGAATACATT	NTGGTCACAT	CATGTAACGC	GTAATACAGC	GATTATAGCT	GTTGGTGGTG	300
	GTGCGACAGG	AGATTTTGCA	GGATTTGTAG	CAGCAACACT	ATTAAGAGGT	GTCCATTTTA	360
40	TACAAGTTCC	TACAACGATT	TTGGCGCATG	ATTCTAGTGT	TGGCGGTAAA	GTGGGTATTA	420
	ACTCAAAACA	AGGTAAAAAC	CTTATCGGTG	CATTTTATCG	TCCAAC TGCT	GTGATTTATG	480
	ATTTAGACTT	TTTAAAGACG	TTACCATTTG	AGCAAATATT	AAGTGGCTAT	GCAGAAGTTT	540
	ATAAGCATGC	GTTATTGAAT	GGTGAATCAA	CGACGCAAGA	AATCGAACAG	CACTTTAAAG	600
45	ATAGAGAGAT	ATTACAGCCA	TTAAATGGTA	TTGGATAAAT	ATATTGCTAA	AGGT	654

## (2) INFORMATION FOR SEQ ID NO:235:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 880 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:235:

CTAACTTTCG	CTTTTCTTT	TTCTTCAATG	TCTTGCACTT	GTCTTCAAC	AATGTTAAAT	60
CTTTGTTGTA	AACCTGCTTT	TTCTGTTTCT	AGTTGTTTGA	TGTCTTCCAT	ATCAATATTT	120
GGATCTGTTG	CTTCTGACT	CAATTCATCA	TTTTTATTTT	TTAATTGNTG	TCCAATCATA	180
CCTAAGGATT	GTTTNAATTC	ATATAATGTC	GGCATNTCAT	TTCTCCTAA	TAANTCATTG	240
TCATTTTAA	AATTTGCGAT	TCGAGCGTAC	AATNTNNTCT	CTNTNNTCTT	NCTCTTCTAG	300
CGACATACTT	TCTTTAGGTG	GTTTCAACCA	ATCCAGATCG	TATCTNACAT	CATCAATTTT	360
AGTGATTTN	TCTACATCTT	TCTTTAAATC	TTCTGGGACN	TTCTCNAAC	GCCTACATTG	420
CTCTTTAGAG	ATACTAGCAG	CTATTTTCATT	AGCTCCTAAA	ATTTTCATCTA	TCAAGCCGAA	480
AGACAAGGCT	TCTTCTGCAG	TAAGCCAAGT	TTCTGCATCT	AACATCTGTT	TTAAGTGTTT	540
TTGATCTAAA	TCTNTTGNNT	TATCTAAATA	AGCTGAATTA	CTAACAGCAT	CTGTTTTTTC	600
AAGTAAATCC	GCTGTCTTTC	TTAATTCTTC	TGCATTACCT	ACAGTCATAA	CCCATGAATT	660
ATGAATCATT	AAAAAACTAT	TNNTGGGGCA	TAAAAATAGT	GTCACCACTC	ATAGGGATAA	720
CAACTAGCAA	TTGGATGCCG	CTAAGGCATC	GACATAGATA	TTANTTNNTG	GAGGATGCAT	780
TTNTAGCATA	TTGGNTTTT	GGATGNTCCN	CAAATACANT	GGCTCCAGAT	GAATTTATAT	840
TGAACAACTT	TTTCAACTGA	TGTCTCCNNG	GTCAACTTAG			880

(2) INFORMATION FOR SEQ ID NO:236:

30

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 830 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: Genomic DNA

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:236:

AGTATTTTTT	TGACCGAGAC	ATGCAAGCCC	TCCTGGTTGA	TTATCATTTA	CATATCGTAT	60
CATAACCAAT	CATAACAAT	AGTTTATTA	ACTTGTAGAC	ACAATGTTTG	CTAAAGTCAT	120
TTTTAAAAA	TATAGCCAAA	AAATTAGCTA	TATATTATAA	AAGCGTGATA	TAAATGTTTT	180
ATATAACAAA	GAAATAAAAA	TCATTTTTTA	CAAATGGTTG	TAAGAAAAAG	ACATGCAGAT	240
GTTGTTAAAA	TTTAAATAAG	AATCAAGGAG	GCTATATTAT	ATGGCTAAAC	TAAATGTAGA	300
AGTATTGCG	GACGGTGCAG	ATATTGAAGA	AATGAAAGCA	GCTTATAAAA	ACAAACAAGT	360
GGATGGTTTT	ACAACAAATC	CTAGTTTAAT	GCCGAAAGCG	GCCGTAAGCA	GATTACAAAG	420
CTTTTGCTGA	AGAAGCTCGT	GAAAGAAATT	CCAGATGCTT	CAATTTTCATT	TGAAGTATTT	480
GCAGACGATT	TAGAACTAT	GGAAAAAGAA	GCAGCAATTT	TAAAAACAATA	TGGCGAAAAAT	540
GTATTTGTTA	AAATTCCTAT	TGTAAATACA	AAAGGTGAAT	CAACGATTCC	TTTAATTAAA	600
AAACTTTCAG	CTGACAAATG	GAGATTAAAC	GTTNCGGCTG	TTTACACAAT	TGAACAAGGT	660



# EP 0 841 394 A2

AAAGAAATAA CTGAAGCAGT AACTGAAGGT GTGCCCAACA TATGTTTCAG TATTTGCAGG 720  
ACGTATTGCA GATACAGGCG TAGATCCATT ACCATTAATG AAAGAGGCTG TAAAAGTTAC 780  
GCATAGTAAA GACGGCGTTA AATTATTATG GGCAAGTTGC CGCTCGTGCC 830

## (2) INFORMATION FOR SEQ ID NO:237:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 789 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:237:

NTTAATTTTA AAATGTACAA TGATATNTGT GAAAGCGCTT GCTTAGGAGG TGTATNTGAG 60  
AGTGAATGAA ATGAATGCTA AAGAACNATT AGTGGACNAT TTAATGAAAA CATCATCGCA 120  
ATTATTTAAA TTNCACGGTG AAGTTGNCAT GCAGCTTNTC TTAAATGATG AATTAAAAATT 180  
ACCTTCTATT GNTGAAATAT GCGTGGAACG TAAGCGTTTA AGTGATATTG TGAAAGTTAT 240  
TCCGCAATCA TATGCGTTAC TATACATAGA TAAGCAAGAT CAAGCAAGAG CTAAAGANNA 300  
TTTATCACTT NCAAAAATTG CAAAAGTTTA TGTGCAATAT GATGATACAA CAATAATGAG 360  
TATTTTCGTT TATGATGTAG TAAACGATGA ATGGATTTTA AGATTTGGATC CGAATATACG 420  
TATACCTAAG AGTAACATAT ACTTCCATAG TTTAAATTGG GATGTGGATT ATATTAAACC 480  
GGAGATCGTT CTAATGTATG ATCTAATGCA ACACCATCAG TATCATCATT ATTCCAATTA 540  
TAAACGAGTC ATAGATGCNT TAAGCTACTA TCAATTTTTT ATTTTAAAAAT TTGTAGTAGG 600  
TGAGCNACGT ATTAAGGATG CAATCCAGAA GAACAATAAA TAATTAAGAA AAAGCAATTC 660  
ATAACGCAGT TGAATACATG TGTTACGAAT TGCTTTTATA TTAGTTTTTA TCACACAAGT 720  
TTTTTAATGC AACNCCGTGA TAGCAAACT CATATGTAGA TAATACAGCT TTTTCAGCAT 780  
CATCTACAT 789

## (2) INFORMATION FOR SEQ ID NO:238:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 605 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:238:

ATATTTCGAGA CGACCCTAAN TAATTGTGTT GTCTGTCATA CTGGGNCANG ATTTGTGTGN 60

EP 0 841 394 A2

CTANATACNA CTNATCATT GCANTAGATT GTTTATATGT TGTGAGTGG CAAAATTGTT 120  
 GTGGTGCGTG TGGGGTTGCN TGCTGAGATG AGTGAGTATA TTGTTGAGTA TTATATGGCT 180  
 5 CTGGTGTTGG GTGTTGCGCC TGTGGCACGT TAGNCTCGAG CCGCTGGGTG TGTATTATCA 240  
 TCAGTTTTCT TCTGAGTATC ATCTGAGNTA TCTTGAGATG CATTGTCATC CTTATCTTTC 300  
 GACTCATCCT NTGATGCTTT ATCATCATTC TCNTCTTTAG CAGGACGTG CTTGTAGGT 360  
 GCTTGTTGCT GAGGTGGCAC ATATTGATAT TGATTCTGCT GTGGCAATG CTGATATTGG 420  
 10 TTTGCTTGTG AATCTAGCTC AGCCTGCTTT TTTTCCTCTT GTTGCTTTNN CTTCTCTCT 480  
 TTATCTGCAA TTTCTTTTGT ACGCTTTTCT TTTTGTTCCT GCTTTTTGTT CGTTCAACAT 540  
 ACGTTCTTTA GCTTTATTCG AATGAATCTA CATATGCAAA AATTGCAAAT ACTAAATCCT 600  
 CCCGA 605

(2) INFORMATION FOR SEQ ID NO:239:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 683 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:239:

GGCACGAGAC TGGTGGAGGT TCAGGTAACCT CGCCGGGCACG ATTAATGCCT GATAAAATAT 60  
 TGGATTAAAA GTATAAGTTG CGTGTTAATA ATGTACCGAC ACCAAGAACA GTAACATTTA 120  
 ATGATACATT AACATATAAA ACATATACAC AAGATTTTAT TAATTCACCT GCTGAAAGTC 180  
 35 ATACTGTAAG TACAAATCCA TATACAATCG ATATCATCAT GAATAAAGAT GCATTACAAG 240  
 CCGAGGTTGA TAGACGCATG GCAACAAGCT GATTATACAT TTGCATCATT AGATATCTTT 300  
 AATGATCTTA AAAGACGTGC ACNAACGATT TTAGATGAAA ATCGTAACAA TGTACCATTA 360  
 AATAAAAGAG TTTCTCAAGC ANATATTGAT TCATTAACTA ATCAAATGCA ACATACGTTA 420  
 40 ATTCGAAGTG TTGATGCTGA AAATGCNGTT AATAAAAAAG TTGACACAAA TGGGAAGATTT 480  
 AGTTAATCAA AATGATGANT TGACAGATGA AGAAAAACAA GCNGCAATAC AAGTTATCGA 540  
 GGAACATAAA AATGAAATAA TTGGTAATAT TGGTGACCAA ACGACTGATG ATGGCGTTAC 600  
 TAGAATCACA AGATCAAGGT ATACAGACCT TAAGTGGGGA TACTGCAACA CCGGTTGTTA 660  
 45 AACCAAATGC TAGGAAAAGC AAT 683

(2) INFORMATION FOR SEQ ID NO:240:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 932 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

# EP 0 841 394 A2

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:240:

```

TGATGATTAG ATAAATTGAA ACAACAACGT ATTGAGTTGA ATGAACAAAT CGATGCGCAG      60
GAAGCTACAT ACAAGTTTGT CACCAAGATA TTTTAGCTAT CGAAAATCAC TACCAAGATA      120
TTAAAGCTGA ACAATCAAAG TTAGGATGTA TTAATTCATC ATGCGATAGA TCATTTAAAT      180
GATGAATATC AATTGACTGT TGAACGTGCG AAATCTGAAT ATACGAGTGA TGAATCGATT      240
GACGCATTAC GTAAAAAAGT TAAGTTAATG AAGATGTCCG ATTGATGAAC TAGGTCCTGT      300
AAACTTAAAT GCAATTGAAC AATTGTGAAG GTTAAATGAA CGTTATACAT TTTTAAAGTGA      360
ACAACGTACA GATCTTCGTA AAGCTAAAGA AACATTAGAG CAAATTATAA GTGAAATGGA      420
TCAAGAGGTT ACTGAAAGAT TTAAAGAAAC TTTCCATGCT ATTCAAGGAC ATTTTACAGC      480
TCGTGCCCAA ACAATTGTTT GGTGGAGGCG ATGCAGAATT GCAATTAACT GAAGCCGATT      540
ATTTAACAGC TGGTATTGAT ATTGTGGTAC AACCACCGGG TAAAAAGTTG CAACATTTAT      600
CGTTACTGAG TGGTGGTGAG CGTGCAATTAA CTGCTATTGC TTTACTATTT GCAATTTTAA      660
AAGTAAGATC TGCACCTTTT GTTATATTAG ATGAGGTTGA AGCTGCACTA GATGAAGCAA      720
ATGTTATTAG ATACGCAAAA TATTTAAATG AGTTATCAGA CGAAACACAA TTCATTGTTA      780
TTACACACCG TAAAGGAACA ATGGAATTG CAGATAGGTT ATACGGTGTA ACAATGCAAG      840
AATCAGGTGT TACTAACTT GTGAGTGTGA ATTTAAATAC AATAGATGAT GTGTTGAAGG      900
AGGAGCAATA ATGAGCGGCA CGAGCTCGTG CC                                  932

```

(2) INFORMATION FOR SEQ ID NO:241:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 966 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:241:

```

CGGTTAATCG GTAGTATTTT AGGTTCTNGTT GTTATTAAGA TTTGGAAGA TTTGATCTTC      60
CATCCCCCAA AGGNTTAAC CGTATGGGGA TNCNTCANAA TCATTGCCC AAGCTAAACC      120
ACCAAAGAAT GGATATTGTA ATCCGGTTGC ATTTAAGNTC AATTGGGATG ATTTTGTAGCC      180
TCCAGTTATT AAGTTCAAGT GTTAAATACC CTGTTCTAAT TTTAATTAAT TCTATTAAAT      240
TAAGTACTTT AAAATGATAA GACACTAAGA AAGGAGGCT ATTAGTAATA ATGCCCAAAA      300
ATAAAAGCAA ATGAAGCATT AGTTAAAGCA TTACAAGCAT GGGATATAGA TCACTTGTAT      360
GGTATTCCAG GAGACTCAAT CGACGCAGTT GTCGATTAGT TTAACGTACA GTGAGAGATC      420
AATTTAAATT TTATCATGTA CGTCATGAAG AAGTAGCAAG CTTAGCGGCT GCTGGTTACA      480
CAAAATTAAC TGGTNAAATC GGTGTGGCAT TAAGTATCGG TGNCCTGGT TTAATTCAAT      540
TATTAAATGG TATGTACGAT GCCAAAATGG ATAATGTACT CGTGCCAATT AATATTATCT      600
GGACAAACNG AATAGTACAG CACTTGAAC GAAAGCATTT CAAGAAACAA ATTTACAAAA      660

```

EP 0 841 394 A2

ATTATGTGAA GATGTAGCCG TTTATAATCA CCAAATTGAA AAAGGTGACA ATGTGTTTGA 720  
AATCGTTAAC GAAGCAATTC GTACGGCATA TGAACAAAAA GGTGTCGCTG TTGTTATTTG 780  
5 TCCTAACGAC TTATTAAC TGAAAAATTAA AGATACAACG AATAAACCCAG TAGATACATC 840  
AAGACCAACA GTTGATATCAC CAAAATATAA AGACATCAAA AAAGCGGTTA AACTAATTAA 900  
TAAAAGTAAA AAGCCTGTCA TGTTAATTGG TGTAGGTGCA AAACATGCCA AAGATGAGCT 960  
CGTGCC 966

(2) INFORMATION FOR SEQ ID NO:242:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1396 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:242:

CTGTGAAATT AAGTCGNTAT TACCCGCAAG GATTAAGAAG TTTGAATGGT GGTCGCATGG 60  
CAAGATTGCG ACGTACACCA TTACTTGATG CAATGGAGAT GGCTAATGAG CATATTATGG 120  
TGATTGCCAT GATAGAAGAT GTTGANGGGG TTATGGCCAT TGACGATATA GCTCAAGTCG 180  
AAGGTTTAGA CATGATAGTC GAAGGTGCCG CAGGATTTAT CGCAGTCACT TGGCATAACCA 240  
30 TNGCAAACGA GCGTGATGAT CAAGTAACAT CACATNTTCA ACATATTTNT GNGGTTGTGA 300  
ATGCACATGG TAAACATTTN TGTGCATTAC CACGTGAAGA TGAAGATATT GCAAAATGGC 360  
AGGCACAAGG TGTACAAACA TTTATTTTAG GCACGAGCGG AAAAATATAT CGCCATTTAA 420  
GTGCATCTCT AGCGACGCTC AAACAGAAAG GGGATGATGG CTAATGCGTA TAGTTCAACC 480  
35 TGTTATTGAA CAATTAAAAG CACAATCTCA TCCAGTTTGT CATTATATCT ATGATTTAGT 540  
CGGACTGGAA CATCATTTGC AACATATTAC ATCGTCATTG CCGAGTAATT GTCAAATGTA 600  
CTATGCAATG AAAGCAAATA GTGAACGAAC AATCCTAGAT ACAATTAGTC AGTATGTTGA 660  
AGGATTGCAA GTTGATCTC AAGGTGAAAT AGCAAAAGGT CTTGCTTTTA AACCAGCAAA 720  
40 TCATATTATT TTTGGTGGCC CTGGTAAGAC AGACGAGGAA CTAAGATATG CAGTAAGTGA 780  
AGGTGTTTCA CGTATTCATG TTGAAAGTAT GCATGAATTA CAACGGCTAA ATGCCATCTT 840  
AGAAGATGAA GATAAGACAC AACACATTTT ATTGCGTGT AATTAGCAG GACCATTTC 900  
CAATGCAACG TTGCATATGG CAGGACGCCC AACACAATTT GGTATTTCTG AAGACGAAGT 960  
45 TGATGATGTC ATTGAAGCTG CGCTAGTAAT GCCAAATATT CATCTAGATG GCTTTCATCT 1020  
TCATTCCTAT TTCTAACAAT TTAGACTCGA ATTTACATGT CGATGTAGGG GAAACTTTAT 1080  
TTTAAAAAAG CAAAATCATG GCGTTGAAAA ACATCGATTT CCACTCAAAC ATATCAATCT 1140  
TGGTGGGTGG CATAGGCGGT CAACTATGCA GATTTAACTA GGCCAACNG AGTGGGATAA 1200  
50 TTTTGNNGAA AATTTNAAA CACTTATCGT TGAGCAAGAA ATGGAAGATG TGACATTGAA 1260  
CTTTGAATGT GGGCGCTNTA TTGTGGCACA TTGGGGTTAC TATGTGACAG AAGTGCTAGA 1320  
TATTANGAAA GNGCATGGCG CTTGGTATGC CATTNTAAGA GGAGGTACGC AACACNTAG 1380  
CCTGCCGGNA TCTTGC 1396

EP 0 841 394 A2

(2) INFORMATION FOR SEQ ID NO:243:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3918 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:243:

ACCAAAGCCA	AATACTTGCA	TTAANNCGGG	GTAATTCCAA	CAACTTAAGA	ACNGTGCGAA	60
ACNAGCCTGC	GGAACNACAA	GCGATTACCA	AATGCTTCGG	GACNTTAAAA	TACCACAACA	120
CAAAGAAGCA	TTGTAAGCAC	AAGTAACAAG	TGCAGGACGC	GTATCTGCAG	CANATGGTGT	180
TGAACATACT	GCGACTGAAT	TAAATACTGC	GATGACAGCT	TTAAAACGTG	CCATTGCTGA	240
TAAAGCTGAC	ACACAAGCTA	GTGGTAATTA	TGTCAATGCT	GATGCGAATA	AACGCCAAGC	300
ATATGATGAA	AAAGTGACAG	CTGCAGAACA	TATCGTTAGT	GGTACACCAA	CACCAACGTT	360
AACNCCATCA	GATGTTACAA	ATGCAGCAAC	GCAAGTAACG	AATGCGAAGA	CGCAGTTANN	420
CGGTAATCAT	AATTTAGAAG	TAGCGACNCA	AAATGCTAAC	ACAGCAATTG	ATGGTTTAAC	480
TTCTTTAAAT	GGTCCTCCAA	AAACCCAAAA	CTTAAAGAAC	AAGTGGGTCA	AGCGACNGAC	540
GTTNCCAAAT	GTTCAAACCTG	TTCGTGATAA	TGCACAAACA	TTAAACACTG	CAATGAAAGG	600
TCTACGAGAT	AGCATTCGGA	ATGAAGCAAC	GATTAAAGCA	GGTCAAAACT	ACACAGATGC	660
AAGTCAAAAC	AAACAAAATG	ACTACAACAA	TGCAGTCACT	GCAGCANAAG	CAATCATTGG	720
TCAAACAACT	AGTCCATCAA	TGATTGCGCA	AGAAATTAAT	CAAGCGAAAG	ACCAAGTGAC	780
AGCTAAACAA	CAAGCGTTAA	ACGGTCAAAG	AAAACCTAAG	AACNCGCAA	ACCAAATGCG	840
AAGCAACAAT	TGAATGGCTT	AAGTGACTTA	ACTTAATGCC	CCCCAAAGAT	GNAGCGANAC	900
CNCCAAATCG	AGGTGCAACG	CATGTTAATG	GAAGTAAACA	CCCAAGCCCA	CAATAATGGG	960
GACGGCATT	AAATACAAGC	TATGNCGGNA	CTTGTNAAAT	GGTAATTCAA	AGACTCACAA	1020
TNCGGATTAA	GCAAGGTGTT	AACTTCACCT	GATGCAGATG	AAGCGAAACG	TAATGCATAT	1080
ACAAATGCAG	TGACGCAAGC	TGANCAAATT	TTAAATAAAG	CACAAGGGCC	AAATACTGCA	1140
AAAGACGGTG	TCGAAACTGC	GTTACAAAAT	GTACAACGTG	CTAAAAACGA	ATTGAGCGGT	1200
AATCAAAATG	TTGCGAACGC	TAAGACAAC	GCGAAAAATG	CATTGAATAA	CCTTACATCA	1260
ATTAATAATG	CACACAAAGC	AGCATTGAAA	TCACAAATTG	AAGGTGCGAC	AACAGTTGCA	1320
GGTGTAATC	AAGTGTCTAC	AATGGCATCT	TGAATTAAAT	ACCTGCAATT	GAGCAACTTA	1380
CCAACGTGGT	ATTAATGACG	AAGCAGCTAC	AAAAGCAGCT	CAGAAATATA	CTGAAGCAGA	1440
TAGAGATAAA	CCCANCCTGC	ATACAATGAT	GCTGTAACAG	CAGCTAAAAC	GTTATTAGAT	1500
AAAACAGCTG	GTTCAAATGA	CAATAAAGTA	GCCGTTGAAC	AAGCATTACA	ACGTGTGAAT	1560
ACTGCTAAAA	CAGCATTAATA	TGGTGACGCG	CGATTAAATG	AAGCGAAGAA	CACAGCTAAA	1620
CAACAATTAG	CGACAATGTC	ACATTTAACT	AATGCTCAAA	AAGCAAACCT	AACAGAACAA	1680
ATTGAACGTG	GTACAACTGT	TGCTGGTGTT	CAAGGCATCC	AAGCAAATGC	TGGTACTTTA	1740
AATCAAGCAA	TGAATCAATT	AAGACAAAGT	ATTGCTTCTA	AAGATGCGAC	TAAATCAAGC	1800
GAAGATTATC	AAGACGCGAA	TGCAGATTTA	CAAAATGCAT	ACAATGATGC	GGTAACTAAT	1860
GCTGGAGGTA	TTATTAGTGC	ANCGAATAAC	CCTGAAATGA	ATCCTGATAC	AATTANCCAA	1920

# EP 0 841 394 A2

	AAAGCGAGCC	AAGTGAACAG	TGCGAAGTCT	GCATTGANC	GTGATGAAAA	ATTAGCAGCA	1980
	GCAAAACAAA	CTGCGAAATC	AGATATCGGT	CGTGTGACAG	ACTTGAACAA	TGCACAACGA	2040
5	ACTGCGNCAA	ATGCTGAAGT	GGATCAAGCA	CCAANTCTTG	CAGCTGTCAC	AGCGGCTAAA	2100
	AATAAAGCAA	CATCGTTAAA	CACAGCGATG	GGTAATGTGA	AACATGCACT	TGCTGAAAAG	2160
	GATAATACGA	NACGTAGTGT	CAATTACACA	GATGCGGATC	AACCAANACA	ACAAGCGTNT	2220
	GATACTGCAG	GTACACAAGC	AGAAGCAATT	ACTAATGCAA	ATGGCAGTNA	CGCGAATGAA	2280
10	ACACAAGTTC	AAGCAGCGCT	TAACCAATTG	AATCAAGCTA	AAAACGACTT	GGAATGGGTG	2340
	ATAATAAAGT	TGCTCAAGCG	AAAGAAACAA	CAAAACGTGC	ATTAGCTTCA	TATAGTAACT	2400
	TGAATAACGC	GCAATCAACT	GCAGCAACTA	GTCAAATTGA	CAATGCAACG	ACAGTAGCAG	2460
	ACGTAAGTGC	TGCACAAAAT	ACTGCTAATG	AATTAAATAC	AGCAATGGGT	CAACTTCAAA	2520
15	ATGGTATTAA	TGACCAAAAC	ACTGTTAAAC	AACAAGTGAA	CTTTACAGAT	GCTGACCAAG	2580
	GTAAGAAAGA	TGCTTACACA	AATGCTGTTA	CGAATGCTCA	AGGTATTTTA	GATAAAGCAA	2640
	CACGGTCCAA	AATATGNCAA	AAGCACAAGT	TGAAGCTGCA	TTAAATCNAG	TANCGNCTGC	2700
	TAAGAATGCT	TTAAACGGTG	ATGCAAATGT	NAGACAANCA	AAATCAGATG	CGAAAGCAAA	2760
20	CTTAGGTACA	TTAACACACT	TANATAATGC	ACAAAAACAA	GATTTAACAT	CACAAAGCGN	2820
	NGGTGCAACA	ACAGNCAACG	GTGTAAATNG	TGTTAAAACG	AAAGCACNAG	ACTTAGATGG	2880
	TGCAATGCAA	CGATTAGAGT	CAGCAATCCG	AAATAAAGAC	CAAATAAAG	CGAGCGAAAA	2940
	CTACATTGAC	GCAGATCCAA	CTAAGAAAAC	AGCATTTGAT	AATGCCATCA	CACAAGCTGA	3000
25	ATCTTACTTA	AATAAAGATC	ATGGTACGAA	TAAAGATAAG	CAAGCTGTTG	AACAAGCAAT	3060
	TCAAAGTGTA	ACGCTTACTG	AAAATGCTTT	GAACGGTGAC	GCGAACTTAC	AACGCGCTAA	3120
	AACTGAAGCT	ACACAAGCTA	TCGATAACTT	GACACAATTG	AATACACCGC	AAAAAACAGC	3180
	ATTGAAACAA	CAAGTGAATG	CTGCACAACG	CGTATCAGGT	GTAACGTATC	TGAAAAATAG	3240
30	TGCTACATCA	CTTAATAATG	CGATGGATCA	ATTAAAACAA	GCAATTGGTG	ATCATGACAC	3300
	AATTGTAGCT	GGTGGTAATT	ACACTAACGC	AAGTCCTGAT	AAACAAGGTG	CTTACACTGA	3360
	TGCATATAAT	GCTGCGAAGA	ATATCGTAAA	TGGTTCACCT	AATGTGATTA	CAAATGCAGC	3420
	AGATGTTACT	GCGNCAACAC	AACGTGTCAA	TAATGCTGAA	ACAAGTTTAA	ATGGTGAGAC	3480
35	AACTTAGCA	ACTGGCGAAG	CAACAAGCTA	AAGATGCATT	ACGTCAAATG	ACACATTTAT	3540
	CTGATGCACA	ANAAACAAAG	TATTACTGGT	CAAATTGATA	GCGCGACACA	AGTAACTGGT	3600
	GTACAAAGTG	TGAAAGACAA	TGCAACAANT	CTTGACAATG	CAATGAATCA	ACTTCGAAAT	3660
	AGTATTGCGA	ATAAAGATGA	AGTAAAAGCG	AGTCAACCAT	ATGTTGATGC	AGATACAGAT	3720
	AAACAAAATG	CATACAATAC	AGCAGTTACA	AGTGCTGAAA	ATATCATTTA	TGCAACGAGT	3780
40	CAGCCAACAC	TTGATCCATC	TGCAGTAACA	CAAGCAGCTA	ATCAAGTGAA	CACTAACAAA	3840
	ACTGCGCTTA	ATGGTGCGCA	AAACTTAGCA	NATAAAAAGC	AAGAAACAAC	TGCTAACATC	3900
	ACCCGATTAA	GTCATTTA					3918

45 (2) INFORMATION FOR SEQ ID NO:244:

## (i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 1265 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

55 (ii) MOLECULE TYPE: Genomic DNA

EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:244:

5	ATGCGATACG	CTTAACCAGC	AATGCATGGT	TTAAGACAGA	GCATTCAAGA	TAACGCAGCA	60
	NCTAAACCGA	ATAGCAAATA	TATCAACGGA	AGTCCACGAG	AGCAACCAAA	ACTATGATCA	120
	AGCTGTTCAA	GCCCCAAATA	ATATTATCAA	TGAACAAACT	GCAACATTAG	ATAATAATGC	180
10	GATTAATCAA	GTAGCGGCAA	CTGTGAATAC	AACGAAAGCA	GCATTACATG	GTGATGTGAA	240
	ATTACAAAAT	GATAAAGATC	ATGCTAAACA	AACGGTTAGC	CAATTAGCAC	ATCTAAACAA	300
	TGCACAAAAA	CATATGGAAG	ATACGTTAAT	TGATAGTGAA	ACAAC TAGAA	CAGCAGTTAA	360
	GCAAGATTTG	ACTGAAGTAC	AAGCATTAGA	TCAACTTATG	GATGCATTAC	AACAAAGTAT	420
15	TGCTGACAAA	GATGCAACAC	GTGCGAGCAG	TGCATATGTC	AATGCAGAAC	CGAATAAAAA	480
	ACAAGCCTAT	GATGAAGCAG	TTCAAAATGC	TGAGTCTATC	ATTGCAGGAT	TAAATAATCC	540
	AACTATCAAT	AAAGGTAATG	TATCAAGTGC	GA CTCAAGCA	GTAATATCAT	CTAAAAATGC	600
	ATTAGATGGT	GTTGAACGAT	TAGCTCAAGA	TAAGCAAAC T	GCTGGAAATT	CTCTAAATCA	660
20	TTTAGATCAA	TTAACACCAG	CTCAACAACA	AGCGCTAGAA	AATCAAATTA	ATAATGCAAC	720
	AACTCGTGAT	AAAGTGGCTG	AAATCATTGC	ACAAGCGCAA	GCATTAAATG	AAGCGATGAA	780
	AGCATTAAAA	NAAAGTATTA	AGGATCAACC	ACAAACTGAA	GCAAGTAGTA	AATTTATTAA	840
	CGAGGATCAA	GCGCAAAAAG	ATGCATATAC	GCAAGCAGTA	CAACACGCTC	GAAGATGCCT	900
25	TGNTTAAACA	AACAAC TGAT	CCTACATTAG	CTAAATCAAT	CATTGATCAA	GCGACACAGG	960
	CAGTGA CTGA	TGCTAAAAAC	AATTTACATG	GTGGATCAAA	AACTAGCTCA	AGATNAGCAA	1020
	CGTGCAACAG	AAACGTTAAA	TAAC TTGTCT	NAC TTGAATA	CACCACAACG	TCAAGCACTT	1080
	GANAA TCAAC	TCAATTCNTG	CAGCAACTCG	NCGNGANGTA	GCACANAAAT	TTACTGANGC	1140
30	ACANGCACNT	TATCCACGCA	ATGGANGCTT	TNCGTCATAG	CATTACAGNN	NCAACAAACA	1200
	AAACAGATCT	GGTNTGCAAN	TTTTTTATTG	AAGGTTNMNC	CNCACCCCGG	TGCTTCCNGC	1260
	GCCAC						1265

(2) INFORMATION FOR SEQ ID NO:245:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 791 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:245:

50	TTCATTAGTG	GTGTGGCAG	ATAGCTCAAN	GATAACCAAA	CTCCTGGAAA	TTCTCTAAAT	60
	TCATTTAGAT	CAATTAACAC	CAGCTCAACA	ACAAGCGCTA	GAAAATCAAA	TTAATAATGC	120
	AACAAC TCGT	GATAAAGTGG	CTGAAATCAT	TGCACAAGCG	CAAGCATTAA	ATGAAGCGAT	180
	GAAAGCATT A	AAAGAAAGTA	TTAAGGATCA	ACCACAAACT	GAAGCAAGTA	GTAAATTTAT	240
	TAACGAGGAT	CAAGCGCAAA	AAGATGCATA	TACGCAAGCA	GTACAACACG	AGCGAAAGAT	300
55	TTGATTAA CA	AAACAAC TGA	TCCTACATTA	GCTAAATCAA	TCATTGATCA	AGCGACACAG	360

EP 0 841 394 A2

	GCAGTGACTG ATGCTAAAA CAATTTACAT GGTGATCAAA AACTAGCTCA AGATAAGCAA	420
	CGTGCAACAG AAACGTTAAA TAACTTGTCT AACTTGAATA CACCACAACG TCAAGCACTT	480
5	GAAAATCAAA TCAATAATGC AGCAACTCGT GGTGAAGTAG CACAAAAATT AACTGAAGCA	540
	CAAGCACTTA ACCAAGCAAT GGAAGCTTTA CGTAATAGCA TTCAAGATCA ACAACAAACA	600
	GAATCTGGTA GCAAGTTTAT TAATGAAGAT AAACCGCAAA AAGATGCTTA CCANGCAGCA	660
	GTTCAANATG CAAAAGATTT AATTAACCAA ACAGGTANTC CAACGCTTGA TAAAGCACAA	720
10	GTTGAACAAT TGACACATGC TTTTAAACAA GCTAAAGATA ACCTACACGG TGATCAAAAA	780
	CTTGACAGCG A	791

(2) INFORMATION FOR SEQ ID NO:246:

15	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 3323 base pairs
	(B) TYPE: nucleic acid
20	(C) STRANDEDNESS: single
	(D) TOPOLOGY: linear
	(ii) MOLECULE TYPE: Genomic DNA

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:246:

	CTATTAAGGA ATCCTTGNN AATTCACATT AGCAAGTTGG ATTGTCCACC TTATTAATTG	60
	ATAAAGTATG CAAATAATTT CACAGTGT TT CCAGCGGGGC CAANTTATGA TTGCAAGTAA	120
30	CCTATAGCAA TCGTATTCTT GTTCTTGCCA ACGTATTTA AGTATCAGGT TTAACCAACA	180
	GGTGGGACAA AAGGTTAGTT TGAAATTAGG GAGTGGGGCA GAATTGATAA AGAACCACTA	240
	ATGACGATAA AGATTAAAGG GGAGGACGTT ATGATGACGG ATTAAAGTTG GAATCATTGG	300
	GTGTGGTGGT ATTGCGAATG GCAAGCAACA TGCCAAGTTT ACNAAAAGTT GAAAATGTTG	360
35	AAATGATCGC ATTTTGTGAC GTAGACATTT CGAAAGCAGC GAGTGGGGCA GAAGCATACG	420
	GAAC TGACAAAGGTT TATGATGATT ACAAAGCATT GTTAAAAGAT GACACGATTG	480
	ATGTTATCCA TGTTTGTACG CCAAATGACT CGCATTGTGA AATTACTGTA GCAGGGTTGC	540
	ATGCCGGTAA GCATGTGATG TGTGAAAAAC CAATGGCTAA AACGACAGCA GAAGCTCAAA	600
40	AAATGATAGA TACAGCTAAA TCAACAGGTA AAAAATTAAC AATAGGTTAT CAAAATCGCT	660
	TCAGACCAGA TAGTCAATTT TTACATCAAG CAGCGCAACG TGGCGACTTA GGAGACATTT	720
	ACTTCGGAAA GGCACATGCC ATTCGTCGTC GTGCAGTACC GACATGGGGT GTCTTTCTAA	780
	ACGAAGAAGC CCAAGGTGGA GGACCTTTAA TCGATATTGG TACGCACGCA TTAGATTTAA	840
45	CGTTATGGAT GATGGATAAT TATGAACCAG AATCAGTGAT GGGTTCAACA TTCCATAAAT	900
	TAAATAAGCA GCATGATGCG CCAAACGCTT GGGGTTTCATG GAATCCAGAT GAATTAACAG	960
	TTGAAGACTC TCGGTTTGGC TTTATTAAAT GAAGAACGGA GCGACGATCA TTTTAGATCC	1020
	GCTTGGGCGA TTAATTCCTTT AGAAGTGGAT GAGGCAAAAT GTTCATTATT AGGAACGAAA	1080
50	GCAGGTGCTG ATATGAAAGA TGTTCTACGT ATTCATGGTG AAGATATGGG CACACTTTAT	1140
	ACCAAACACG TTGAATTNGA AAACAAAGGC GTCGACTTTT ATGAAGGTAA TGAAGTGGAT	1200
	GAAGCTGAAG AAGAANCAAA AGCTTGATTG GATGCAGTTG TAAATGATAC TGAACCAATT	1260
	GTGAAACCGG AACAAGCAAT GGTAGTTACA AAAATTCTTG AAGCGATTTA TCAGTCTGCA	1320
55	AAATCAGGCA AAGCAATTTA CTTTGAATAA CATCATACGG TAAGGAGGCA CATCATGACA	1380



# EP 0 841 394 A2

	AAATTAAAAG	TTGGTGTGAT	AGGTGTTGGT	GGCATTGCAC	AAGACCGTCA	TATTCCAGCA	1440
	TTGCTGAAAC	TCAAAGACAC	AGTCTCATTA	GTTGCAGTAC	AAGATATTAA	TACAGTGCAG	1500
5	ATGATTGATG	TTGCGAAGCG	CTTTAATATA	CCTCAGGCAG	TTGAGACACC	TAGCGAGCTG	1560
	TTTAAACTTG	TTGATGCGGT	GGTCATTTGT	ACACCCAATA	AATTCCATGC	TGATCTTTCT	1620
	ATAGAAGCAT	TGAACCATGG	TGTCCATGTC	TTATGCGAAA	AGCCAATGNC	AATGACGACG	1680
	GAAGAGTGTG	ATCGCATGAT	TGAAGCGGCT	AATAAAAAATC	ACAAATTATT	AACTGTCGCT	1740
10	TATCATTATC	GTCACACAGA	TGTCGCTATG	ACTGCTAAAA	AAGCAATTGA	AGCAGGTGTG	1800
	GTTGGTAAAC	CATTAGTAGC	ACCGTTGTCC	AANCGATGCG	TAGGCGTAAA	GNACCTGGGT	1860
	GGGCGTTTTT	TACCCAATAA	AACCGTTGCA	AGGTGGCGGT	AGTTTAATCG	ATTATGGTTT	1920
	CCCAC TTGTT	AAGACTTATC	TTTGTGGCTA	TTAGGTAAAG	ATATGGTGCC	GCATGAAGTG	1980
15	CTAGAAAAAA	CATATAATCA	ATTGAGCAAA	CAACCGAATC	AAATTAATGA	TTGGGGAACA	2040
	TTTGATCACA	CTAAATTTGA	TGTCGATGAT	CATGTTACTA	GTTATATGAC	ATTTGCCAAT	2100
	CGAGCAAGCA	TGCAGTTTGA	ATGTTCTGGG	TCTGCAATAA	TAAAGGAAGA	TAAGGTACAC	2160
	GTTAGTTTAT	CAGGAGAAGA	TGGCGGTATC	AATTTATTTC	CATTTGAAAT	ATATGAGCCC	2220
20	CGCTTTGGAA	CTATTTTGA	AAGCAAAGCT	AATGTTGAGC	ATAACGAAGA	CATTGCTGGT	2280
	GAGAGACAGG	CGCGTAACTT	TGTCAATGCG	TGTTTAGGTA	TAGAAGAGAT	TGTGGTGAAA	2340
	CCGGAAGAAG	CACGCAATGT	AAATGCCCTT	ATAGAAGCGA	TTTATCGTAG	CGATCTTGAT	2400
	AACAAGAGCA	TACAAC TTTA	ATGATTATTA	TATATAATAC	AAAAT TCTCA	ATATAAAAAG	2460
25	ANGGAGTGCT	TTTCAATGAA	AATAGGTGTA	TTTTCAGTAT	TATTTTACGA	TAAAAATTTT	2520
	GAAGATATGT	TAGATTATGT	CGCAGAATCT	GGATTGGATA	TGATTGAAGT	TGGAACAGGT	2580
	GGTAACCCAG	GAGATAAATT	TTGTAAGTTA	GATGAGTTGT	TAGAAAATGA	AGACAAGCGA	2640
	CAAGCATTTA	TGAAGTCAAT	CACAGACAGA	GGCTTACAAA	TAAGTGGTTT	CAGTTGTCAT	2700
30	AACAATCCAA	TTTCTCCAGA	TCCGATAGAA	GCGAAAGAAG	CCGATGAAAC	GTTACGTAAA	2760
	ACAAATCCGTT	TAGCAAATCT	ATTAGACGTG	CCAGTTGTTA	ATACATTTTC	TGNCATTGCA	2820
	GGGTCAGATG	ATACCGCTAA	AAAGCCTAAT	TGGCCTGTTA	CACCTTGGCC	AACAGCCTAC	2880
	TCTGAAATTT	ATGATTATCA	GTGGAATGAA	AAGTTGATAC	CATATTGGCA	AGATTTAGCT	2940
35	GAGTTTCCNC	AAGAGCAAGA	TGTNNAAATT	GCCATNGAGT	TACATNCAGG	ATTTTTAGCT	3000
	CANACACCAA	ATACGATGTT	NAAGTTACGT	GAGCCAACAA	ATGAATATAT	CGGTGCTAAC	3060
	TTAGATCCTA	GTCATTTATG	GTGNCAAGGT	ATTGACCCAA	NTCCTGCGAN	TCGCATATTA	3120
	GGCCCNANCA	AATNCAATTG	ATCACTTCCA	TGGAATTCCG	AAGAAACGTA	TGTTANTCNA	3180
	GGGNATGTAA	ANATGTATGG	TCTAGCTGAT	NTCCAGCCAT	ATGGTNACGT	TGCGACANGN	3240
40	GCATGGACAT	TCCNTACAGN	TGGNTATGGA	CATAGTCCAT	ATGNATGGGC	AGATNTCATA	3300
	AGTCAACTTA	NTATTAGATG	GAG				3323

## (2) INFORMATION FOR SEQ ID NO:247:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 641 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

# EP 0 841 394 A2

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:247:

5	TCCATTTAAT TCCACCTTGC CCCGACAGGT TTNCCGACC TGGAAAGCGG TCAGTGAGCC	60
	CNAACNCAAT TAATGTGAGT TAGCTCCACT TCATTAGGCC ACCCCAGGCC TTTACACTTT	120
	ATGCTTCCGG CTCCGTATGT TGTGTGGAAT TTGTGAGCGG ATAACAATTT CACACAGGAA	180
	ACAGCTATGN CCATGATTAC NCCAAGCTCG AAATTAACCC TCACTAAAGG GAACACAAGC	240
10	TGGAGCTCCC ACCGCGGTGG CGGCCGCTCT AGAACTAGTG GATCCCCCGG GCTGCAGGAA	300
	TTCGACGAGA ACGCGTGCAG CGTTTACAGT TGCCTCTATT GATTTAGGTG CGCATCCAGA	360
	ATTTTTAGGG AAAAATGATA TTCAATTAGN CAAAAAAGAA TCTGTAGAGG ATACTNCNAA	420
	AGTATTAGGT AGAATGTTTC ATGGTATTGA ATTCAAACTT TAACTGAGCT TCATGGGGCA	480
15	CCAGGTTTTC AAGAAGAAGT AAAAAATTAT ATGACTCAGC AAATGGCGCC GTATGTAGAT	540
	GGAATTTATT GAAAATCGTA TGGGTGGATT TTTTGGGTGT GAAAAATCT AAAAAATCAA	600
	ATGCAAAACG TGTAATGATT GCAGNACATA TGGATGAAAT C	641

## 20 (2) INFORMATION FOR SEQ ID NO:248:

### (i) SEQUENCE CHARACTERISTICS:

	(A) LENGTH: 1130 base pairs
25	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: single
	(D) TOPOLOGY: linear

### 30 (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:248:

35	CTGCTTAAAT TCGAACGTGT CATTGACATT TTCAATAAGA AAAACAACGG AGCATGGAAT	60
	TTCTGTCCGC AAAACGCCCG TTATTGGGAA CATATCCCTA AGAGTATTAC AAAATTATCA	120
	GATTTAAAAA TCGTGGTTT AGATTCTAT ATCACTACTG AAGAATCAAA ACGATTACT	180
	GATTTTCCTA AAGACTTTAA AGGTATTGCA GGTTGGATAT TAGAAGTAAA ATCGAATACA	240
40	CCAGGTAACA CAACACAAGT ATTAAGACGT AATAACTTCC CGTCTGCACA TCAATTTNTA	300
	GTAGAAACT TTGGCACGAG ACTGGTGGCG TTGGTAAATG GGAGTTTATT CGAAGGAAAG	360
	GTGGTTGAAT AATGATAGTA GATAATTTT CAAAAGACGA TAACTTAATC GAGTTACAAA	420
	CAACATCACA ATATAATCCA ATTATTGACA CAAACATCAG TTTCTATGAA TCAGATAGAG	480
45	GAAGTGGTGT TTTAAATTTT GCAGTAACTA AGAATAACAG ACCGTTATCT ATAAGTTCTG	540
	AACATGTTAA GACATCTATC GTGTAAAAAC CCGATGATTA TAACGTAGAT AGAGGCGCTT	600
	ATATTACAGA CGAATTAACG ATAGTAGACG CAATTAATGG GCGTTTGACG TATGTGATAC	660
	CGAATGAATT TTTAAAACAT TCAGGCAAGG TGCATGCTCA GGCATTCTTT ACACAAAACG	720
50	GGAGTAATAA TGTGTTGTT GAACGTCAAT TTAGCTTCAA TATTGAAAAT GATTTAGTTA	780
	GTGGGTNTGA TGGTATAACA AAGCTTGTTT ATATCAAATC TATTCAAGAT ACTATCGAAG	840
	CTGTCGGTAA AGACTTTAAC CAATTAAAGC AAAATATGGC TGATACACAA ACGTTAATAG	900
	CAAAAGTGAA TGATAGTGCG ACAAAGGCA TTCAACAAAT CGAAATCAAG CAAAACGAAG	960
55	CTATACAAGC TATTACTGCG ACGCAAACTA GTGCAACACA AGCTGTTACA GCTGAAGTCG	1020
	GATAAANTAG GTGGGAAAAA GAGCAAGCGA TTGGGGACCG TTTTAACGAA GTNGGACAAC	1080

AAATCAATGC GCCTGGCCTT GTNAAGGTAA TNCAACAACA AATTGGGAAA 1130

5 (2) INFORMATION FOR SEQ ID NO:249:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1104 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:249:

20	CCTCAGCATA TGGTTGAACC CCTTCTTGAT AACCCATATC TGTTAGTTTT CCCCAATGAT	60
	TGCTCGCATA GCTAATTGTG CCATTGGAAT ATCAAGTAAT TTTACTTAGG CACGAGAATG	120
	GTNCCGTTCT ACTAGAACGG TGGGTTTACT TCTAAAACAA TATACACCAT CGTGAGCTAT	180
	AACGAATTGA ATGTTGATTA AGCCAATGAT GTTTAAACCT TTAGCTAATT TTATAGTATA	240
25	GTCCTCAAGT GTTGCTAACT CGTCTTCTGT CAAAGTTTGT GCGGATATA CAGCGATTGA	300
	GTCACCACTA TGTACACCAG CACGTTCAAT ATGTTCCATG ATTCCTGGAA TAATGACCGT	360
	TTCTCCATCA CAAATCGCAT CAACTTCAAT TTCTTTACCA GTTAAATATC TATCGACTAG	420
	TACCGGATGT TCCGGACTCG CTTTACAGC CTGGGTCATA TAGTTTCTA ACTCTTTGTC	480
30	ATTGTCTACA ATTTCCATTG CGCGACCACC TAATACATAA GAAGGTCTTA CTACAACCGG	540
	ATATCCGATT TCTGCAGCAT TCGCTAATGC TTCCTCAGGT GATGTAGCTG APTTCCCTTG	600
	TGGCTGTGGC ACGTTAATTT TTCTTAATAG TGCTCAAAT TCTTTTCTAT CTTCAGCACG	660
	ATTTAGATTT TCTAGTGAAG TACCAAGTAT TTTAACACCA TGTTTAGCCA ATTTGTCTGC	720
35	TAAATTAATC GCTGTTTGTC CTCCAAATTG TACAACGACA CCTTTAGGTT TTTCTAAATT	780
	AATGATATTC ATCACATCTT CTTCAAGTTAA AGGTTCAAAG TATAATTTGT CAGAAATTGA	840
	AGAAGTCTGT TGAAACTGTT TCTGGATTGT TATTCACAAT TATCGCTTCG TACCCTGCTT	900
	TTTGAATTGC CCAAACGGCG TGAAGTGTG CATAGTCAAA TTCTACACCT TGGCCGATTC	960
40	GAATTGGTCC AGAGCCTAAT ACTAAGATT CTCTTTGTCA GTAACATGG ATTCATTTTC	1020
	AGTTTCGTAT GTACCATAAT AATATGGTGT TGAGATCAAA TCAGCTGCGC AGTATCAACA	1080
	TCTGTAACAG GTTAATATCA TTTC	1104

45 (2) INFORMATION FOR SEQ ID NO:250:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1995 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

55 (ii) MOLECULE TYPE: Genomic DNA

EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:250:

5	CTATCATAAA AACATAAAGT ATTGTAAGCT TTTTATCGAT ATTTTATTATT TATAAAAATA	60
	AAATGAGATA ACTNTGTGAA TTTTATTGGA GATAAATTAG ATAGTGNTGT TTTGTGATG	120
	TTTAATAATA TCTTGGGTGT GTTAATACTA GTAATGCTTT CAACTGATGC ATTAAGACTG	180
	TGACATCATA ACTCATTAA GAACCTCGCT TATTAATTTT CTACCAATAC AATCCCTTCT	240
10	AAGTGCACTA AAAAATCCTT ACTGCTAAGT GATTAACTT AACAATAAGG ATTTATTTAT	300
	CGCTAATGCA TGATTATTAA CGGAATCTCA TACCACCATC TACAATAATT GTTGTGCCAG	360
	TAATGTAATC AGAGTCTTTA CCAGCTAAGA AGCTCACTAC ATTTGAAACA TCTTCTGGTT	420
	GAGAACTCT GCCCAAAGCA ATCTGACTTG TAAATTGTTT CCAACCCCAT GCTTCAGGTT	480
15	TACCTGCTTC TTCGGCTGTT GCCACTGCGA TACTTTCCAT CATTGGTGTG TGAACGATAC	540
	CAGGTGCGAA TGCATTCACA GTAATACCTT CAGACGCTAA ATCTTGTGCG GCTACTTGTG	600
	TTAAACCTCG CACTGCGAAT TTTGTACTGC AATATAAAGA CAAGCCTGGG TTACCCCTCAA	660
	CGCCTGCTTG AGATGTTGCA TTGATAATTT TACCGCCATG ATTGAATTTT TTAAATTGTT	720
20	CATGTGCGGC TTGAATACCC CATAGCACAC CTGCAACGTT CACGCCATAT ACTGTTTTAA	780
	ACTGTTCTTC AGTAATTGTA TCGATTGGTG TTGTTGGTCC AAGGCCGGCA TTGTTAACCA	840
	TGACATGGAA ATCGCCAAAT TCGCTCGGA GTTGCTGTG TACTGCGTT AAATACATCA	900
	TCACGGTTTG ATACATCTGC TTTGATAGCA ATAGCTTTTG TACCATCACT TGATAATTTA	960
25	AGTGCACTG CTTTTCGCC TTTCTCATTT AAATCAACAA CTGCTACTTT GAAACCATCT	1020
	TCCACTAAAC GTTCTGCAAT TTTAAAACCA ATCCCTGTGT CTCCGCCAGT TACTAATGCT	1080
	ACTTTGTGTT TTGTCATAAA GATCACTCCT CAAATTCTTT TCCTTTAATT ACATTTTACT	1140
	CCTCTTCATT TGAATAGTAC AACAAAGGTA GCTCCATTTA ACAAATATT CAGATATTTA	1200
30	AGGTATAGTT AAACGCACTA CCATTAGTGA TTGGCAATGC GTTTAAATGT CGTTTTAAAA	1260
	GTTCTTATGT TGAATATTAT TTTTAAAGT CTCTCGATTA GTTTGTGATC AATCTTTTTT	1320
	CGAGACATGG GCCTTTTGAT TCAATCGGCG GNTTCCGTGT TATCACTGAC AACTTTAGTT	1380
	GTAGCTTCAT CTTTATGTAT TTCTTCGCTA AATCCTTCAA GGTTTTTAGT CGTGGGATTT	1440
35	TTAACCTCAG GATGTTCCAT CATGCTTTTG ACTATCAAGT TCCTTTTAC ACGTGTCTTT	1500
	ATGGTGATGC TTGATTGGC GTTCCCTTTT ACTTTTTTTG AATAGTGGG AGGTATCTGC	1560
	TGCAGCTACT AATTNCTTC TTACCCTTAA AATAGATATT GGCTTGGAAC CAAAACCAGA	1620
	GTATTTGCAG ATACAAAGTT GCATTAATAC TTAAAGCAAT AATAGCCAAT ACAAATTAAT	1680
40	ATTGACACCT NTTGAAATCC CTTCTTTTAA ATTAAGTCAG ATGCCAATAC GATGACAGGG	1740
	TACGGATTGA AAGTATAATT ACAAATATAG AAATTATTGC CGATATAACT ATTGTTACTA	1800
	TTAAATAATC AGCTCTGCTA CCTGATAATA AATAGAAAAG GCGAAAATTA GTCCATAGCA	1860
	AATTACAAAC CCACATAAAG TTATAGCCAT GAGTACTATA TAAGCTATTT GAAAATATAA	1920
45	ACCTATCTTT ATGAATGATT TTCTACATTT TTTCCATGTC TATCCCCATT ATTAATATTA	1980
	TACTTACCTA ATATA	1995

(2) INFORMATION FOR SEQ ID NO:251:

50	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 1940 base pairs
	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: single
55	(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:251:

	TGTACTTGAA	CACTGTACTG	CGGATGAGTT	TTTTATTGAC	ACGGAATTAT	TAGGGTTCAA	60
10	CGGTGACTTA	CCACGTTTCG	TAATTACCTC	CAAACGACAG	CTAGTACGTT	TATAAGATGG	120
	TGTATCCGTG	TATTGATCAA	CATCACTATT	AGTTAATAAG	TTAATTGCAC	CTAGATCTCC	180
	ATTTTCCATC	GCATCATTAT	TTAATGGAAT	ATAGATTTCT	TTACCTTTAA	CACGATCTGT	240
	CACGTGAACT	TGTAATACCG	CTTCTCCTGT	TTCAGAAATC	AGCTTAACTT	CTGCACCTTC	300
15	ATGAATGCCT	CTATCTTCAG	CAAGCTCTGG	AGAAATTTCA	ACAAATGCAC	GTGGCACTTT	360
	GTATTTAATC	ATTGGTGTTT	GATAAGTCAT	ATTACCTTCA	TGGAAGTGCT	CTAACAATCG	420
	ACCATTGTTT	ACATGAATAT	CATAAATTTT	ATCTTGCTTA	AAGTAATTAT	CAAATGATAA	480
	TGGGAATAAT	TTTGCTTTAC	CATTATCAAA	ATTGAATCCT	TCTAAGTATA	GAATAGGCTC	540
20	ATCAGTNCCA	TCCNGNTTGT	ACTTGNCCAT	TGTAAACTAT	TGAATCCTTC	TANNCGATTC	600
	ATAACTTACC	CCAGCATATA	GAGGTGTTAA	GCGTCCCTACC	TTCAATCCAT	AATTTCACTA	660
	GGATGCTTGT	AATTCCAATC	CAAATCCCTA	ATCTATTAGC	AATTGCTTGG	AAAATTTTCC	720
	AGTCAGGTTT	TGAATCACCA	AGAGGTTCTA	ATGCTTGGTA	TAAACGTTGA	ATACGACGTT	780
25	CGGTATTTGT	AAAAGTACCG	TCTTTTTCAA	GTGAAGGACT	TGCTGGCAAT	ACAACATCTG	840
	CGTATGTTGC	TGTGAATGTT	AAAAATTCAT	CTTGGACTAC	CATGAAATCT	AATTTTTCAA	900
	ACGCAGCTTG	TACAAAATTA	ATATTTGAAT	CCACAATACC	CGTATCTTCA	CCATATAAGT	960
	ACAATGAGTG	TACTTCTCCG	TCATGTATAC	CTTCTACCAT	TTCATGATTA	TCTTTACCAG	1020
30	CTTTTGGAAT	CAATTTAACG	CCATATTCTT	TTTCAAATTT	AGCACGAGCG	AATATCATCC	1080
	GCTTCAATAC	TTTGATAAACC	AGTAATCTTA	TCAGGCATAC	TTCCCATATC	ACTACATCCT	1140
	TGAACATTAT	TATGTCCACG	TAATGGATAC	GCACCAGTAC	CAGGACGACG	ATAATTACCT	1200
	GTTACTAATA	ATAAGTTTGA	AATCGCTGTA	CTTGAGTCAC	TACCAATGTC	TTGTTGTGTA	1260
35	ATACCCATTG	CCCAACAAAT	TACAACAGAT	TCAGCTTTAG	CACATTCTTC	AGCAAATTTA	1320
	ATCAATTCTG	ATTGAGGAAT	ACCTGTTGCT	TCTTCAGCAA	AAGCCATTGT	AAATGTTTCT	1380
	AATGATTTGT	AATATTCATC	AAAATCATCT	ACCCACTCAT	CAATAAATGC	TTTATCGTGT	1440
	AAATCATGAT	CAATAATATA	CTTAGTCACT	GCACTTAACC	ACGCTAAATC	CGTACCTGGT	1500
40	TTAGGTTGAT	AAAAACGATC	CGCACGTTCT	GCCATTTTCA	GTTTTCTAAT	ATCAAATACA	1560
	TGTATTTTTT	GTCCAAATAA	TTTTTGAGCA	CGTTTCATGC	GTGATGCGAT	AACTGGATGA	1620
	GCTTCGGCTG	TATTGGTACC	TATCAATACA	GACATTGCCG	CTTTTCTTAA	ATCTTCAATA	1680
	CTACCTGAGT	CACCGCCGTG	TCCAACCGTT	CTAAATAAGC	CTTTTGTGTC	AGGTGCTTGG	1740
45	CAATATCTTG	AACAGNTATC	AACGTTATTT	GTGNCAATAA	CTTGGTCTTG	CTTAATTTNN	1800
	GGGATGNAAA	TACGATTCTT	CCATTCCGGC	CGGTTTTAGA	AGAAGGAAAT	GAATTGATAG	1860
	TGCATCTTGG	GCCAATACTT	NNNCTTTNAA	TAGATGTAAA	ATTATCTGCA	AATGACAATT	1920
	TAAAGGTTCC	ATCCCATCTT					1940

(2) INFORMATION FOR SEQ ID NO:252:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 847 base pairs

(B) TYPE: nucleic acid

# EP 0 841 394 A2

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:252:

```

10  GATTCAACAT AACGTCGTGT CCTTCAGCAT ATATAGTATC GAATGCTATG GACGATTAAC      60
    CTGACCCAGA TAAACCTGTC ATAACAATTA ATTTATTTT AGGTAGTTCG GATATCAATA      120
    TCTTTCAAGT TATGCGCAGC AGCACCTTTT ACTACTATGG ATGGNTCTTT CATTTACTTG      180
    TCACCCNTCT GCTTTTAATT CAAATAACAT ATCTCTTAAT TCNGTAGCTT TCTCGAAATC      240
15  TAAATCTTTC GCTGCTTGTT TCATTCTTTT TTCTATATG TCGATTGTCT TTNGACGTGC      300
    TCTTTTCGGC ATCTTCTTAG GTATCACAGT TTGTGCTTNG TCATNATTTT CGTCATTTTC      360
    AACAGTAGCA CTAATTAAAT CATGTACTTN TTTATTAATT GTTTTAGGTG TAATACCATG      420
    TTTTTCATTA TGTTCATCT GTATTTCTCG ACGACGTTGT GTCTCATCAA TTGCATACTT      480
20  CATNGAATCA GTCATATTAT CGGCATACAT AATGANTTCA CATTTATCGN TACGCTCGTG      540
    CCGAATCGGC ACGAGCTCGT ACCTATTGTT TGAATTAATG AGCGGNTAGA ACGTAAAAAT      600
    CCTTCTTTAT CTGCATCTAA TATGACAAAGT AGAGAACTT CTGGTATATC AATACCCCTCT      660
    ACTTAATAAA TTAATACCTA CGATAACATC ATATGTACCC ATTCGTAAGT CTCTAATTAT      720
25  TTCGATTCGT TCGTAATGTC TTGATTTCTG AATGCAGATA ATTAACTTTA ATACCCGCTT      780
    CTTTCATGTA TGTGGTTAAA TCTTCACTCA TCTTTTtagT GAGCGTTGTA ACAAGTACGC      840
    TCGTGCC                                         847

```

30

(2) INFORMATION FOR SEQ ID NO:253:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1300 base pairs

35

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

40

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:253:

```

45  CCGCTTCGCA CGAGCNAAGG CTTTATTGTC ANCCCTCGCA TCTTCTTtag CTTACACTAA      60
    ATCTTTGATG TCTTGTAATG CTTGTGCAAG TTGTGCTTTG GCTTGTCAA TTTCTTCTTT      120
    AGTCATCGCA TTGTTAATGT CGTTATGACC TTGTTGAAGT ATTTGANTAA TTCGATCTTT      180
    AAGCGCTTGT TTNTCCTTAT CTGTTAGATT TGGATTGTGA TCGATTTCGN CAATTAATGC      240
50  TTGAACCTGT TTATCAACAT CTTGTTTCGC ATCTTCTTTA GCTTACACTA AATCTTTAAT      300
    GTCTTGTAAT GCTTGCGCAA GTTGTGCTTT AGCTTGTTCA ATTTCTTCTT TAGTCAGCGC      360
    ATTGTTAATG TCGTTATGAC CTTGTTGAAG TATTTGATTA ATACGATCTT TTAAGTGCTT      420
    GTTTTTCCTT ATCTGTTAGA TTTGATTTTC GATCAATTTT GTCAATTAAT GCTTGAACCT      480
55  GCTTATCAAC GTCTTTATTG GCATCAATTT TNGCTTTNGG TATTTCATTA GCANGCACTN      540

```

# EP 0 841 394 A2

	GTTCAATCGC GTGGTTGCCT GCTGTTTGAA CTTGAGATAC AGCCTGATTA CTTGTTGCTT	600
	TATTAATGTT GTTGATGATG CTGGTTTGCC AATTCTTCTT GCTTTATTTT TTCGGCAATA	660
5	AGCTTGTCCT GATCCGTCGC ATTTGAAGCT TCGATTCTT TTAGCTTATT AGCTAAAGCT	720
	TGATTAATTG ATTGAATTGC CTTGTCTTTA GCATCTTGTA GTCGTTGATC ACCATTAAGA	780
	TTATGGATTG CATCATTGAC TGCTTGATT GCGCCATTGA TATCATTCAC ATTTGTGTTA	840
	TCACTATTTA GCAATGTATT TGCTAGACGT TTGGCATCAT CGAAGTTTGT TTTAGCATT	900
10	TCGTCAGCGT TTTGGTAATT GACAGTTTGC TCTGCATTG GAATTTTCATT GTCAACTAAA	960
	TGTTTCAATG TTCCATTGC ATCATTTAAG TCAATTTGAT TATTAACAAT ATCTGTTACA	1020
	TCTGATACAG TATCGGCATT GTTAATTGCT TTATGTGCAA GATCTTGTG CTGTTGATTT	1080
	AATCCATTTA ACGAATTAAC AAACGCATTT GCTTTATCCT TTGGCATTG CAAGGTTTTG	1140
15	GTCTCCATTT AATGCATTTT GAGCATCGAT AATATTTTGT TTCAATTGCT CTGCTTCAGC	1200
	TTTTGCAATT GCATTACTTG CACTTTATCT AACTCATGTG CTGCATTGCG AATGCATATC	1260
	ATAATTTGCT TTCAATGTCA TCTGCATGAT GTATTGTGCT	1300

20 (2) INFORMATION FOR SEQ ID NO:254:

(i) SEQUENCE CHARACTERISTICS:

25 (A) LENGTH: 1425 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:254:

35	CTTTTCATCT CTTGACCTG TGTAATACCT TTGATCAGCT GCACAAATAG ACAAGTATAA	60
	AGGTCTTACG ATGACATGAT TACCATAAAT ATCAACATTA TTATATGTGA CGTCGAACTG	120
	TCTCGGTGCA ACGAGTTGAT ATACTTGATT AATCATCGGC AATATCACCT TGAATAATGG	180
	NATTTGCTAC TTTNAANTCA TNCGGGGTTG TCACTTTAAT GTTGTATAGT TCTCCACGTA	240
40	CCAATTTAAC TGCATGTCCA GATTCGACAA TTATTTTACA TGCATCTGAT AAGATTTCTT	300
	TNTGGTTTAC TACTTAAGGC GCGATAACTA TCTTGTAATA ATTTAATATT AAATGATTGT	360
	GGTGTTTGGC CTTGATACAT TTCATTCCCT ACAGGGATAC TGTGTATGTT CTGTTTATCT	420
	TTAGACATTA CAATCGTATC AATTGCTTCA ATGACTGTAT CTACTGCACC ATATTNNGCT	480
45	GCTACTTCAA TGTTCTCTTT AATAATACGT TGAGTTAAAA ATGGTCTTAC GGCATCATGA	540
	GTTACAATCA CATCATCATT ATTAATTCCA TTTACATTGC GAATATGGTC GATAATGTTT	600
	ATAATTGTTT CGTTTCGATC CGNACCACCT GCAACTACTT TGACACGTTG ATCTGTAATG	660
	TTATATTTTT TTAAAAATATC CTGTGTATGG GAAATCCACT GTGCTGGCGT TGCGATAATA	720
50	ATCTCATTAAT ATTCATTCAC TAAAAATGAAC TTCTCAATTG TATGGATTAA AATCGGTTTA	780
	TTATCAATAT CTAAAAATTG TTTTGGGTAA TGGCACGTTT CCCATTCTCG AACCAATACC	840
	TCCTGCTAAA ATACCTGCAT AAATCATGTT GGCCTCCATC CTGTCATTAC ATCATTTCCT	900
	TTTATACATT ACTGACCTAT GCCCGCACAT AAGCCTAACC TATTGCTCAC TNGNCTCTTN	960
55	TATTAATCCA AAGATAGTTG TCACAATAGT GTGATAATTN TTTATAAAAA TGTATTTNTG	1020
	NNACTGACCA NTCTAAGTTG TTTTGGCATG CAGCTAATCA TTAACCTCTGA CGATATTAAA	1080

EP 0 841 394 A2

TIGTTAAAGG TATTAATGTT TACTCTTTCA CAAATTCAATT ATTACTGCCA TCATTTNNQC 1140  
 ATATATTATA ATAAATTTAT CTTATTAAGT GGCTGNACTT GATTTTCACT TTAAAAATTA 1200  
 5 TCAAATATTG CCATCTCAT TTAAGNATAC AAAATGCAAA ACAACCGATT CACAAGCATA 1260  
 TTTCACACAA GTAAACCGGC TATTTATCAA CGTATATTCG AAGATGAATT ATTTTCGATAG 1320  
 TATCTTGTAG ACCAGNCGGC ATTCGCACTT TCAATAGCNT ATTAACCTAT ACCAGNGGTT 1380  
 TCGTCCTCNA NGGTGCATAC TAATAAATCG TAAACNTGAC TTTAG 1425

(2) INFORMATION FOR SEQ ID NO:255:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3160 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:255:

TTTCTTACCT TTAATAATTG GTAATTCATC TTCCATCCTC TTGTTTGATC TGTGTGCGCAT 60  
 AAGCTTCATC AAATAGTACA TCAGTACGTG GAACACCTGT AGGCACTACA TTTTCTCTTT 120  
 TAATACCAAA TGCTTCAGCG TAGAATGGAA TATCGGTTTC AGATGATACA TAAGCTTTGG 180  
 30 CACGAGCTAC GATGATTTAA TGAATCAATA AATGGTCCAC CCTTTTACC AGTACGACTA 240  
 AAGCCAACTG TTTTAAAGGC ACCAACGGCA TGCCATACCT GAATAACTTC TTGAGAACGT 300  
 CTAAAACGCA CTGTATAAAT CAAGGGGTGA AAGTCATCAA CAAAGATGTA GTCTGCCCTTC 360  
 CCAAGTAAAT ATGGCAATCT AAACCTGTCG ATGATGCCAC GTCTATCTGT AATATTCGCT 420  
 35 TTAAAAACAG TGGAATATC ATACTTTTAA TCTAAATTTT GACGTAACAT TTCGTTATAG 480  
 ATGTATTCAA AGTTCCAGA CATCGTTGGT CTAGAGCTCG ATGTGAACAA CACCGTATTC 540  
 CCTTTTTC AAGGTGAAAA ATTTTCGTCG ATTAATATC GCNTNAAAAA TAAATGTCT 600  
 TTGTTATTAA ATGAATGTTT GCGGAAATAC NTACGTAATT CTTTAATATT TACGAACGAA 660  
 40 AATAAATACT TTAACTTCC GGAGTCGTTA CNACCAACAT CAAGGACAAA TTCATTAAACA 720  
 TTCGCNAGAA ATTTACAGGT GTAACAGTAT AAACCGTTT CTTCGAAATG CCGCCTTTTC 780  
 TAAATCTTT TAGGTAAAGT TGCAATAAGA AATTGATTT ACCATTTTGT GTTCTAATT 840  
 CGTTGTATTC TTCTTCTTGT TCTGGCTTTA GATTTTGATA TGCATCATTA ATCACATCTG 900  
 45 GGTTTANCTG TGCAATATAA TCAAGTTCTT GCTCATTCAC TAATAAGTAC TTATCTTCAG 960  
 GTAAGTAATA ACCATTATCT AAGATAGCTA CATTGAAACG ACAAACGAAT TGATTCACAT 1020  
 CTATTTTGAC ATCATTCGCC TTCATTGTAC GTGTCTCAGT TAAATTTCTT AATACAAAAT 1080  
 TACTATCTTC TAAATCTAGG TTTTCACTAT GTCCTTCAAC GAATAACTGA ACACGTTCCC 1140  
 50 AATAGATTTT ATCTATATAT ATCTTACTTT TANCCAACGT TAATTCATCC TTTTCTATTT 1200  
 ACATAATCCA TTTAATACT GTTTTACCCC AAGATGTAGA CAGGTCTGCT TCAAAAGCTT 1260  
 CTGTAAGATC ATNAATTGTT GCAATTCAT ATTCTTGACC TTTTAACAAC GCTAATTTAT 1320  
 CTACAATATC TGGGTATTGA ATGTATAAGT CTACAACATC TTGGAAATCT TGTGACCCAC 1380  
 55 TTGCACTACT ACCAATCAAC GTTAGTCCTT TTTCCAATAC TAGACGTGTA TTAACCTCTA 1440  
 CTGGGAATC ACTTACACCT AACAGTGCAA TGCTTCCTTC TGGTGAAATA TANTCAATCA 1500



EP 0 841 394 A2

	TTTGATTAT	TGCTGATTGG	NTACCACGCC	CACCAACGCA	CTCAAATGCA	TGATCANATG	1560
	TTAAGCCTTC	AGGTATTTTA	TTAATAAAGA	AGACATCATC	AACAAATGAG	AAGTGACTCA	1620
5	ATTTATAATC	TTGGTNNACC	AAAGACATAT	ATGGNAGACT	TCTGGATATT	AATNGACGGA	1680
	ATAAAATGGC	TGTAATGTAA	CCTAAGTTAC	CATCACCCCA	AATACCAAAT	GTATTTTTAT	1740
	TTGAAATAGA	TTTCTTTTCA	ACAACGACGA	ATAGCATGCA	AACTTACTGT	TACAAGCTCT	1800
	GTAGATGAAA	TAATACTTAA	ATCAATATCA	TCAGGTAGTG	GTACAGCTCT	ATCATGATTT	1860
10	AGCAACACAA	AATCTTGCA	AAACCCATCA	TGTCCACTTG	ATCTGAAGTA	GCTCGATTTT	1920
	AAATAGTTTT	CAGCAATGAC	ATCGTCTTTT	TCTGTCGGCG	TATTCGGTAC	CATAACTACT	1980
	TTTGTACCTT	TATTAAACAC	ACCTTTACTG	TCAAATACGA	CCTCACCAAC	ACCTTCATGA	2040
	ATTAAAGACA	TTGGCAATTT	CTGAGATAAG	ACATTCTCAT	CACGGCTACC	AGTATAATAT	2100
15	CTTTGATCGG	CAGCACAAAT	TGACATATAT	AAAGGACGTA	CAATGACATA	GTCACTGTAA	2160
	ATATCTACGT	TGTTATACGT	AACTTCAAAT	TGTCTAGGTG	CAACTAATG	ATATACTTGA	2220
	TTAATCATCG	GCCAATACCA	CCTCGAATAA	TAGCATTCGC	TACTTTTAAA	TCGTAAGGTG	2280
	TTGTTACTTT	AATGTTATAT	AACTCACCTT	TTACAAGTCG	AACCGGTTTG	TTTGTTCCTA	2340
20	CAATAATCTT	ACAAGCATCA	GATAAAATAC	TCTTTTGCTC	ATCACTCAAC	TGTGCATAGC	2400
	TTTCTTTTAA	TAAATTAATA	TTAAACGATT	GAGGTGTTTG	ACCTTGGTAC	ATTTCATTAC	2460
	GCACTGGAAT	TGCATCAATC	GTTTGATTAT	CTTTAGATGT	AACAATCGTA	TCTATAGCAT	2520
	CAATCACTGT	ATCTACTGCA	CCGTACTCTA	AAGCAGCTTG	AATATTTTCT	TTAATAATAC	2580
25	GATGCGTTAA	AAATGGTCTA	ACTGCATCAT	GTGTCACAAT	GACATCGTCA	TCGTAAATAC	2640
	CATTTGTGTA	TTCAATATGT	TTAACGATAT	TCATAATTGT	ATCGTTACGA	TCGCTACCAC	2700
	CTTGAATGAC	TTCAATTCTT	TCATCAGAAA	TTTTGAATTT	TCTAAGTGTA	TCTTTCGTAT	2760
	GCGTCATCCA	TTGTTGTGGC	GTCGCGATAA	TAATTTTTC	AAAATCATT	ATTAAAATAA	2820
30	ATTTTTCTTA	ATGTATGGAT	TAAATCGGN	TTGTTGTCTA	AATCTAAAAA	TNGTTTAGGT	2880
	AAAGGGTACG	GTACCCATT	CCTTGAGCCT	ATACCTCCAG	CTAGAATACC	AGCGTATTTT	2940
	ATAAAATACT	TCCTCCATT	AACTATATCT	ATATTTAATT	ATTAAATTT	CGTTGCATTT	3000
	TCCAATTGAA	AACTCATTTT	AAAATCAAAA	CTCTAAATGT	CTGTGTATTA	CTTAAAATTA	3060
35	TACATATTTT	GCTTATATTT	TAGCATATTT	TGTTTAAACC	TATATTACAT	TATATCAGAC	3120
	GTTTTCATAC	ACAAATAATA	ACATACAAGC	AAACATTTCT			3160

(2) INFORMATION FOR SEQ ID NO:255:

40	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 1676 base pairs
	(B) TYPE: nucleic acid
45	(C) STRANDEDNESS: single
	(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:256:
----	---

	AAAAAAACAA	TTTANCCCAT	NACNACTTGT	CGTTAATTAT	TCATACGAAA	TACATGATTA	60
55	ATGTACCACT	TTAACATAAC	AAAAATCGT	TATCCATTCA	TAACGTATGT	GTTTACACAT	120
	TTATGAATTA	GATAACGATT	GGATCGATTA	TTTTATTTTA	CAAAATGACA	ATTCAGTTGG	180

EP 0 841 394 A2

	AAGGTGATTG	CTTTTGATTG	AATCGCCTTA	TGCATGAAAA	ATCAAAAGGT	TATTCTCAT	240
	GTATAGTCCT	GCTTCTCATC	ATGACATGTT	GCTCACTTCA	TTGTCAGAAC	CCTTCTTGAA	300
5	AACTATGCCT	TATGACTCAT	TTGCATGGCA	AGTAATATAT	GCCAACATTA	GCGTCTAAAC	360
	AAATCTTNGA	CTAAACGTTT	ACTNGAGCGA	CCATCTGGAT	ATTTAAAANG	TTTAGCTAAG	420
	AATGGTACAA	CTTTTTC AAC	CTCATAATCT	TCAGTGTCCA	AAGCATCCAT	TAATGCATCA	480
	AAGGATTGTA	CAATTTTACC	TGGAACAAAT	GATCCATATG	GTTTCATAGAA	ATCACGCGTC	540
10	GTAATGTAAT	CTTCTAAGTC	AAATGCATAG	AAAATCATCG	GCTTTTTTAAA	TACTGCATAT	600
	TCATATATTA	AAGATGAATA	ATCACTAATC	AACAAGTCTG	TAACAAAGAG	AATATCGTTA	660
	ACTTCACGAT	GATCTGACAC	ATCGATAAAG	TATTGTCTAT	GTTTACGTGA	AATATTAAGT	720
	CTATTTTTTA	CGAACGGATG	NATTTTGAAT	AACACAAC TG	NATTATGCTT	CTCGCAGGAT	780
15	CTTGCTAAAC	GTTCAAAATC	AATTTTAAAA	AATGGATAAT	GTGCCGTACC	GTGACCATTA	840
	CCTCTAAATG	TCGGTGC GAA	TAGAATAACT	TTCTTACCTT	TTATAATTGG	CAATTCATCT	900
	TCCATCTCTT	GTTTAATTTG	TGTTGCATAA	GCTTCATCAA	ATAGTACATC	AGTACGTGGT	960
	ACACCTGTTG	GTACAACATT	TTCTTCTCTA	ATTCCAAATG	CTTCAGCATA	AAATGGAATA	1020
20	TCGGGTTTCT	GATGAAACAT	ATGCTTTCGG	TGTAAC TACG	ATGGGTTTTA	ATGAATCGAT	1080
	AAACGGACCA	CCTTTTTT NAC	CTGTACGACT	AAAGCCA ACT	GTTTTTAAAGG	CACCAACAGC	1140
	ATGNCACACT	TGAATAATT	CTTGTGATGG	TCTAAAGCGA	ACCGTATAAA	TTAATGGATG	1200
	GAAATCATCA	ACAAAAATAT	AATCGGCCTT	ACCAAGTAAA	TATGGCAATC	TAAACTTGTC	1260
25	TCTCCATT TG	CGTCTATCCG	TAATATTCGC	CTTAAATACC	GTTTTAATAT	CATAATCAAA	1320
	ATCTACTTTG	TGGCGTAGTA	ACTCATCATA	TACATACTTG	AAATTCCCTG	ATAAATTCGG	1380
	ACGCGAATCT	GATGTGAATA	ATATTGTTNT	GCTTCTTTTA	ATATGTAGTA	ACTTTGTAAT	1440
	ATTAAAAATA	GCTTTAAATA	AGAAACTTCT	ACTTTCAAAT	GAAGCTTTAT	GACCTTGTTT	1500
30	ATGAAGCCAG	TGTGCACTTG	GCGCAATGAC	CCTGATT TCT	CTTGAGGTAA	GGTGATTTC A	1560
	ATATCAAATA	CAAAATTCGGN	TAACCGTCAC	TTGGCTATCT	CCGGAGTAAT	GGTATTAGAC	1620
	CGTATGTTGT	GATACGCCAC	CTTTACGGAA	AACTTTAGCA	TCATACGCTA	ATAAAG	1676

(2) INFORMATION FOR SEQ ID NO:257:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3110 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:257:

	GATTTCGAGCT	CGGTACCGGG	GATCCTCTAG	AGTCGACCTG	CAAGCTATAC	CAAGCTAGAG	60
50	TACTGCGACG	CAAACTAGTG	CAACACCTAA	GGCTGTTACA	GCTTGGAAGT	CGATAAAAATA	120
	GTTGAAAAAG	AGCAAGCGAT	TTTTGAACGT	GTTAACGAAG	TTGAACAACA	AATCAATGGC	180
	GCTGACCTTG	TTAAAGGTAA	TTCAACAACA	AATTGGCAAA	AGTCTAAACT	TACAGATGAT	240
	TACGGTAAAG	CAATCGAATC	GTCTGAGCAG	TCCATAGATA	GCGTTTTTAA	CACAGTTAAC	300
55	ACATCTAGGA	TTATTCATAT	TACTAATGCA	ACAGATGCGC	CAGAAAAGAC	GGATATAGGC	360

EP 0 841 394 A2

	ACGTTAGAGA	AGCCTGGACA	AGATGGTGTT	GATGACGGTT	CTTCGTTCTGA	TGAATCAAQT	420
	TATACATCAA	GCAAATCTGG	TGTCTTAGTT	GTTTATGTTG	TTGATAATAA	TACTGCTCGG	480
5	TGCAACATGG	TACCCAGACG	ATTCAAACGA	TGAGTTCACA	AAATACAAAN	TCTTACGGCA	540
	CATGGTACCC	GTTTTATTAA	AAAGAATGAT	GGAAACTTAA	CTAGCAATTT	GGTTGGAGGA	600
	ACGCTTTACA	ACGCTTTAAA	TCAAGCTAAG	CAGTATGTAG	ATGATAAAAT	CGGAACAACG	660
	AGCTGGCAAC	AACATAAGAT	GACAGAGCCG	AACGGTCAAT	CAATTCAAGT	TAACTTAAAT	720
10	AATGCGCAAG	GCGATTTGGG	ATATTTAACT	GCTGGTAATT	ACTATGCAAC	AAGAGTGCCG	780
	GATTTACCAG	GTAGCGTTGA	AAGTTATGAG	GGTTATTTAT	CGGTATTCGT	TAAAGATGAT	840
	ACAAACAAGC	TATTTAACTT	CACACCTTAT	AACTCTAAAA	AGATTTACAC	ACGATCAATC	900
	ACAAACGGAA	GACTTGAGCA	ACAGTGGACA	GTTCTTAATG	AACATAAGTC	AACGGTATTG	960
15	TTTCGACGGT	GAGCAAAATG	TGTAGGTACA	ACAATCAATC	TAACTGAACC	GTACACAAAC	1020
	TATTCTATTT	TGTTGGTAAG	TGGAACCTAT	CCAGGTGGCG	TTATTGAGGG	ATTCGGACTA	1080
	ACCGCATTAC	CTAATGCAAT	TCAATTAACC	AAACCGAATG	TAGTTGACTC	AGACGGTANC	1140
	GGTGGCGGTA	TTTATGAGTG	TTTACTATCA	AAAACAAGTA	GTACCCACTT	TAAGAATAGA	1200
20	CAACGATGTG	TATTTTCGATT	TAGGCAAAAC	ATCAGGTTCT	GGAACGAATG	CCAACAAAGT	1260
	TACTATAACT	AAAATTATGG	GGTGGAAATA	ATGAAAATAA	CAGTAAACGA	TAAAAACGAA	1320
	GTTATCGGAT	ACGTTAATAC	TGGCGGTTTA	CGCAATAGTT	TAGATGTAGA	TGATAACAAT	1380
	GTGCCTATCA	AATTCAAAGA	AGAGTTTGAA	CCTAGAAAAGT	TTGTTTTTAC	TAAACGGCGAA	1440
25	ATTAATATA	ACAGCAATTT	TGAAAAAGAA	GACGTACCGA	ATGCATCAAG	CCAACAAAGT	1500
	GAATCAGATT	TGAGTGATGA	AGAACTTCGC	GGAATGGTTG	CGAGTATGCA	AATGCAGGTG	1560
	GCACAAGTAA	ACGTATTAAAC	AATGGAATTA	GCTCAACAAA	ACGCTATGTT	AACACAACAG	1620
	TTGACTGAAC	TGAAAACATA	CAAAACAAGT	ACTGAGGGGG	ACGTTTAAATG	ATGAAGATGA	1680
30	TTTATCCGAC	TTTTAAAGAC	ATCAAAACTT	TTTATGTTTG	GGGTACTAT	AAAAACGAGC	1740
	AAATTAAGTG	GTACGTAGAC	AAGGGTTTAA	TCGATAAAGA	AGAATACGCT	TTAATCACTG	1800
	GAGAAAAATA	TCCAGAAACA	AAAGATGAAA	AGTCACAGGT	GTAATGCTTG	TGGCTTTTTTA	1860
	ATTTAACAAA	AAGTAGGTGG	CGTAATGTTT	GGTTTACCA	AACGACATGA	ACAAGATTGG	1920
35	CGTTTAAACG	GATTAGAAGA	AAATGATAAG	ACTATGTTTG	AAAAATTCGA	CAGAATAGAA	1980
	GATAGTCTTA	GAGCGCAAGA	AAAGATTTAT	GACAAATTAG	ATAGAAATTT	TGGAAGAATT	2040
	AAAGCGCGAC	AAGGTAGAAG	ATGAAAAGAA	TAAAGAAAAG	AATGCCAAGA	ATATTAGAGA	2100
	CATAAAAATG	TGGATTCTCG	GTTTGATAGG	GACTATCTTC	AGTACGATTG	TCATAGCTTT	2160
40	ACTAAGAACT	ATTTTTGGTA	TTTAAAGGAG	GTGATTACCA	TGCTTAAAGG	GGATTTTAGG	2220
	ATATAGCTNC	TGGGCGGTCC	TCCTGGTTGG	GGGTAAATGT	AAATAACAGT	TAAAGATCAG	2280
	TGCTCCGGCA	CTGGCTTTTT	ATTTTGATTG	AAATGAGGTG	CATACATGGG	ATTACCTAAT	2340
	CCAAAGACTA	GAAAGCTAC	AGCTAGTGAA	GTGGTGGAGT	GGGCAAAGTC	GAATATTGGT	2400
45	AAGAGGATTA	ATATAGATAA	TTATCGGGGC	AGTCAATGTT	GGGATACACC	TAACTTTATT	2460
	TTTAAAGAT	ATTGGGGTTT	TGTAACATGG	GGCAATGCTA	AGGATATGGC	TAATTACAGA	2520
	TATCCTAAGG	GTTTCCGATT	CTATCGTTAT	TCATCTGGAT	TTGTTCCGGA	ACCCGGAGAC	2580
	ATAGCAGTTT	GGCACCTGG	CAACGGAATA	GGTTCGGACG	GACACACCGC	AATAGTAGTA	2640
50	GGACCATCTA	ATAAAAGTTA	TTTTTATAGC	GTTGACCAAA	ACTGGGTAA	TTCTAATAGT	2700
	TGGACAGGTT	CTCCAGGAAG	ATTAGTAAGA	CACCCTTATG	TAAGTGTTAC	AGGCTTTGTT	2760
	AGGCCTCCGT	ACTCAAAAGA	TACTAGCAAA	CCTAGTAGTA	CTGATACAAG	TTCAGCATCA	2820
	AAAAGCCAAT	GACTCAACAA	TTACTGGCGA	AGCGAAGAAA	CCGCAATTTA	AAGAAGTTAA	2880
55	AACAGTAAAA	TACACTGCTT	ACAGCAATGT	TTAGATAAAG	AAGAGCACTT	CATTGATCAT	2940
	ATAGTTGTAA	TGGGTGATGA	ACGCTCAGAT	ATTCAAGGAT	TATATATAAA	AGAATCAATG	3000

# EP 0 841 394 A2

CATATGCGTT CTGTAGACGA ACTTTATACG CAAAGAAATA AGTTTATAAG CGATTATGAA 3060  
ATACCGCATT TATATGTCGA TAGAGAGGCT ACAGGAATTC CGGAATTCGG 3110

(2) INFORMATION FOR SEQ ID NO:258:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2553 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:258:

AAATTATATA TAATAACGTC GTTTTAACTA AGGCAACATA AGGAGGTGCG TCAATTGGCT 60  
CAAGCAAAAA AGAAATCGAC AGCTAAGAAA AAAACAGCAT CAAAAAAG AACAAATTCA 120  
AGGAAAAAGA AGAATGATAA TCCGATACGT TATGTCATAG CTATTTTAGT AGTTGTATTA 180  
ATGGTGTGG GTGTTTTCCA ATTAGGAATA ATCGGTCGTC TAATTGACAG CTTCTTTAAT 240  
TATTTATTTG GGTACAGTAG ATATTTAACA TATATTTTAG TACTCTTAGC AACTGGTTTT 300  
ATTACATACT CTAAACGTAT TCCTAAACT AGACGAACGG CTGGTTCGAT TGTATTGCAA 360  
ATTGCATTGC TATTTGTATC ACAGTTAGTT TTTCATTTTA ATAGTGGTAT CAAAGCTGAA 420  
AGAGAACCTG TACTTCTTA TGTATATCAG TCATACCAAC ACAGTCATTT TCCAAATTTT 480  
GGTGGCGGTG TATTAGTTT TTATTTATTA GAGTTAAGCG TACCTTTAAT TTCATTATTT 540  
GGTGATGTA TTATTACTAT TTTATTATTA TGCTCAAGTG TTATTTTATT AACAAACCAT 600  
CAACATCGTG ATGTTGCAAA AGTTGCACTG GAAAATATAA AAGCTTGGTT TGGTTCATTT 660  
AATGAAAAAA TGTCGGAAAG AAACCAAGAA AAACAATTGA AGCGTGAAGA AAAAGCGAGA 720  
CTTAAAGAAG AACAAAAGGC ACGTCAAAAT GAACAGCCAC AAATAAAGA TGTGAGTGAT 780  
TTTACGGAAG TGCCTCAAGA AAGAGATATT CCAATTTATG GGCATACTGA AAATGAAAGT 840  
AAAAGCCAGT GTCAACCAAG TCGAAAAAAA CGAGTGTGTTG ATGCAGAGAA TAGTTCGAAT 900  
AACATCGTAA ATCATCAAGC AGATCAGCAA GAGCAATTAA CAGAACAAAC TCATAACAGT 960  
GTTGAAAGTG AAAACACTAT TGAAGAAGCT GGTGAAGTTA CGAATGTATC GTATGTTGTT 1020  
CCACCGTTAA CTTTACTTAA TCAACCTGCA AAACAAAAAG CAACATCTAA AGCTGAAGTA 1080  
CAACGTAAAG GACAAGTACC AAAGAGATAC ATNAAAAGAT TTTGGGGTNA AATCCNAAAG 1140  
TNGACACAAA TNAAATTGTC CTNCAGTAAC TCAATATGAA ATCCCACCCA GCTCAANGGG 1200  
GGTTNAAAGT GAGTAAATTT GTAACTTTG CATAATGATA TTGCATTAGC TTTAGCACCA 1260  
AAAGATGTTA GAATCGAAGC ACCAATACCT GGTGCTCTG CAGTAGGTAT TGAAGTGCCA 1320  
AATGAGAAAA TTTCAATAGT TTTACTAAAA GAAGTTTATG ATGAAAAATT CCCGTCTAAT 1380  
AATAAACTAG AAGTTGGATT AGGAAGAGAT ATATCAGGTG ATCCAATTAC TGTTCCACTA 1440  
AATGAAATGC CACACTTATT GGTGCCAGGA TCGACGGTA GTGGTAAATC TGTTTGATA 1500  
AATGGTATTA TTACAAGTAT TTTATTAAAT GCTAAGCCGC ATGAAGTTAA ACTTATGTTA 1560  
ATCGATCCGA AAATGGTTGA ACTAAATGTT TATAACGNNN NNNCACATTT ATTAATTCCG 1620  
GTTGTTACAA ATCCTCATAA AGCTGCTCAA GCTTTAGAAA AAATTGTAGC TGAGATGGAA 1680  
AGACGTTATG ATTTATTCCA ACATTCATCA ACTAGAAATA TTAAAGGTTA TAACGAATTA 1740

EP 0 841 394 A2

	ATCCGTAAGC	AAAATCAAGA	ATTAGATGAG	AAGCAACAG	AATTACCTTA	TATCGTTGTT	1800
	ATTGTAGATG	AGCTTGCAGA	TTTAATGATG	GTAGCTGGTA	AAGAAGTTGA	AAATGCGATT	1860
5	CAACGTATCA	CACAAATGGC	ACGTGCAGCA	GGTATACATT	TGATTGTAGC	AACACAAAGA	1920
	CCTTCTGTGG	ATGTAATTAC	AGGTATCATT	AAAAATAACA	TTCCATCTAG	AATTGCTTTT	1980
	GCTGTGAGTT	CTCAACACAGA	TTCAAGAACT	ATTATTGGTA	CTGGCGGCGC	AGAAAAGTTA	2040
	CTTGGTAAAG	GGTGACATGT	TATACGTTGG	AAATGGTGAT	TCATCACAAA	CACGTATTCC	2100
10	AAGGGGCGTT	TTTAAGTGAC	CAAGAGGTGC	AAGATGTTGT	AAATTATGTA	GTAGAACAAC	2160
	AACAGGCAAA	TTATGTAAAA	GAAATGGAAC	CAGATGCACC	AGTGGATAAA	TCGGAAATGA	2220
	AAAGTGAAGA	TGCTTTATAT	GATGAAGCGT	ATTTGTTTGT	TGTTGAACAA	CAAAAGGCAA	2280
	GTACATCATT	GTTACAACGC	CAATTTAGAA	TTGGTTATAA	TAGAGCATCT	AGGTTGATGG	2340
15	ATGATTTAGA	ACGCAATCAG	GTAATCGGTC	CACAAAAAGG	AAGCAAGCCT	AGACAAGTTT	2400
	TAATAGATCT	TAATAATGAC	GAGGTGTAAA	AAATTTGATT	AAGCAAAATA	AATTGGAATA	2460
	TGGTGACCAA	CTATCCAAGT	AATTTATCAA	TTGCCAGAGA	ATTAAATGTA	AAAACCGACG	2520
	ATGTTTATGA	AGCAATTCAG	GCATTGATTA	CTG			2553

(2) INFORMATION FOR SEQ ID NO:259:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1664 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:259:

35	CCGTACATTA	ACAACCTGGA	TTAATTACAG	ATTACGTTTT	GTCTCATCTA	AAAATTGTTG	60
	ATAATCTTTG	ATCAATTTCC	AAATCTGTAC	GTATGTATTA	TCTTAGTAGT	AGTTTAATAT	120
	GCACTAAATA	CTTGCAATTG	ATATTTACCA	TATTTATTGT	CAAATCCAAT	TATCTTGTGT	180
	TTTTCATAAA	ACGATTGCTT	TAAATAATCT	TCTAACACAT	CAAACATCGT	ATTATCACCG	240
40	ACATGGTGCC	CGTATAAAAT	AGTATTATGA	TTTAAATTCT	NCAATTCATT	TCTAAAATCC	300
	ATAAAAATAC	TACCTTTACG	TCGATGTTCT	CGCTCAAAAT	CTAAATTAA	ATAATCGTGA	360
	TTTGTCTTAC	CTTGTAAGTAC	TGGATAANTT	AATGATGTTT	CTGATAATTT	TATCCATCCA	420
	ACAATGTCTT	TATTTATTTT	TTCAAGTGAT	TCAAATTTGT	GTCTCACATG	TTCTTGATGT	480
45	TTGNTCATCA	GCATNTGAAA	TTTNTGTTGT	AATTTCTCAT	AANTTGCGCG	TTCTTGCTTG	540
	TCTTCAATAT	ATGTTTGAAC	AATTTTGTA	CCAAAAATGA	TAATAATTAC	AACCAATAAA	600
	ATTTGTACAA	TAGTTAAAAA	TCGCTTCATT	CTCATAAAAA	CCCTCTTTAA	TAAACAANTT	660
	TCTTCAGTCA	TCACTAAACC	AGTTGTTGTA	CCGTTTGTAG	TTCGATTTCG	TTGACTTTGA	720
50	CAAATTAAGT	AAATTAGCAT	TGGACCACCG	ACAATCATTA	AAATAGCATT	GGCTGGAATT	780
	TCTAAAGGAG	GCTGTATCAC	TCGTCCTAAT	AAATCAGCCA	CTAACAAATAG	CCATGCACCA	840
	ATAACTGTAG	AAAACGGAAT	AAGTACTCTG	TAATTGCCCC	CAACTAGCTT	TCTAACCACA	900
	TGTGGCACAA	TAATACCTAA	AAAGGCTAGT	TGTCCAACAA	TCGCAACAGT	TGCACTTGCT	960
55	AAAAATACTG	CTAATAAACC	TGTTAACCAT	CTGTAACGAT	CAATATTAAA	ACCGATACTT	1020

EP 0 841 394 A2

CGCGCTTGTA TGTCGTCTAA ATTTAGTAAA TTCAATTTAG GGGACAATAG TAATGTTAAT 1080  
 ATTAATCCCA ATAATGCTGA TACTGCTAAT ATGTATACGT CGCTCCATAT TTTCATTGTT 1140  
 5 AAGCCTTGAG GAATTTTCAT TAAAGGGTTT TGAGTTAAAA TTTCTAAAAC ACCATTTAAT 1200  
 AATACGNATA ACGCAACACC TACTAATATC ATACTTACAG CATTGAATCT AAATTTAGAA 1260  
 TGCAACAATA TAATTATTAA AAATGGTATT AAACCTCCAA TAAAACTTAA TAATGGTAAG 1320  
 TAAAAGTACA ATTGTGGAAT AAACAACATA CCAAGTGCTC TCATTATAAG TGCACCTGAG 1380  
 10 GAAACGCCAA TGATATTCGC CTCTGCCAAA GGATTTTGTA GTCCTGCTTG TAATAATGCT 1440  
 CCAGAACTG CTAACATTGC GCCANCCATC AATGCAANTA ATAGACGTGC CAATCCCAA 1500  
 TCAATGATTG AATCCACTGC TTCATTGCTA CCAGTTGTAA ATTTTGTAAG TAGGTCATTA 1560  
 AATGACAATT TAATTGTACC GGTACAAAC GAAATATAAG CAGTTGCGAT TAAAATGACT 1620  
 15 AACAAACATA AAAATAGTAG TTTCTTTTCA TTTTATCA TCAG 1664

(2) INFORMATION FOR SEQ ID NO:260:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:260:

Val Leu His Xaa Ile Arg Glu Ile Leu Leu Gln Leu Val His His Ser  
 1 5 10 15  
 Asp Xaa Gly Xaa Gly Ile Gln Pro Pro Lys Phe Phe Gln Pro Gly Asp  
 35 20 25 30  
 Xaa Xaa Lys Val Thr Ile Asp Asn Ile Gly Thr Leu Thr Thr Tyr Ile  
 35 40 45  
 Xaa Lys  
 40 50

(2) INFORMATION FOR SEQ ID NO:261:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 51 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:261:

EP 0 841 394 A2

Val Pro Ile Lys Ala Val Ile Ala Gln Thr Met Thr Thr Leu Pro Arg  
 1 5 10 15  
 Val Ile Asn Ile Lys Arg Leu Thr Ser Gly Leu Asn Lys Pro Leu Ser  
 5 20 25 30  
 Phe Met Met Pro Lys Lys Ile Met Ala Xaa Ile Asn Ile Ile Thr Thr  
 35 40 45  
 Xaa Met Xaa  
 10 50

(2) INFORMATION FOR SEQ ID NO:262:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 90 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:262:

Val Ile Lys Phe Lys Asn Val Thr Lys Arg Tyr Gly Lys His Val Ala  
 1 5 10 15  
 Val Asp Asn Ile Ser Phe Asn Ile Asn Glu Gly Glu Phe Phe Val Leu  
 20 25 30  
 Ile Gly Pro Ser Gly Cys Gly Lys Thr Thr Thr Leu Lys Met Ile Asn  
 35 40 45  
 Arg Leu Ile His Leu Ser Glu Gly Tyr Ile Tyr Phe Lys Asp Lys Pro  
 50 55 60  
 Ile Ser Asp Tyr Pro Val Leu Arg Asn Ala Leu Gly Tyr Trp Xaa Arg  
 40 65 70 75 80  
 Xaa Gly Xaa Arg Leu Glu Leu Xaa Pro Ile  
 85 90

(2) INFORMATION FOR SEQ ID NO:263:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 90 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

# EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:263:

```

5      Val Cys Cys Phe Gly Ala Ser Thr Met Leu Ser His Ile Thr Ala Tyr
      1          5          10          15
      Glu Ile Leu Ser Glu Ile Xaa Lys Lys Leu Ala Gln Lys Leu Met Arg
          20          25          30
10     Leu Pro Leu Xaa Val Val Glu Ser Lys Lys Ile Gly Glu Leu Lys Asn
          35          40          45
      Ile Phe Xaa Asp Lys Val Glu Thr Ile Glu Leu Pro Leu Ala His Met
          50          55          60
15     Ile Pro Glu Val Xaa Gly Asn Leu Leu Val Ala Ala Ala Ile Phe Leu
      65          70          75          80
      Tyr Ile Met Xaa Xaa Asp Trp Arg Ile Gly
          85          90

```

20

(2) INFORMATION FOR SEQ ID NO:264:

(i) SEQUENCE CHARACTERISTICS:

25

- (A) LENGTH: 152 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:264:

35

```

      Val Thr Ser Phe Ile Tyr Lys Ile Leu Tyr Val Val Lys Ile Asn Ala
      1          5          10          15
      Tyr Thr Tyr Asp Ile Met Thr Glu Asp Ile Met Xaa Leu Ser Ile Leu
          20          25          30
40     Leu Ile Phe Leu Cys Ile Arg Leu Val Ser Leu Lys Ile Cys Ile Xaa
          35          40          45
      His Ser Lys Gln Leu Lys Ala Asp Gly Ala Val Glu Tyr Gly Val Lys
          50          55          60
45     Asn Ser Lys Phe Leu Ala Ile Thr His Val Leu Ile Tyr Val Leu Ala
      65          70          75          80
      Gly Val Glu Ala Phe Ile Asn Lys Asp Thr Phe Ser Phe Ala Asn Gly
          85          90          95
50     Ile Gly Leu Val Ile Leu Ile Phe Ala Tyr Ile Met Leu Phe Met Val
          100          105          110
      Ile Lys Thr Leu Gly Gly Ile Trp Thr Leu Lys Leu Phe Ile Leu Pro
          115          120          125
55

```



EP 0 841 394 A2

.. Asn His Pro Ile Ile Lys Ser Gly Leu Tyr Lys Ile Thr Lys Thr Pro ..  
 130 135 140  
 5 Lys Leu Leu Leu Lys His His Ser  
 145 150

(2) INFORMATION FOR SEQ ID NO:265:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 104 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:265:

Val Arg Val Thr Gly Ile Xaa Pro Tyr Gly Ala Phe Val Glu Thr Pro  
 1 5 10 15  
 25 Asn His Thr Glu Gly Leu Ile His Ile Ser Glu Ile Met Asp Asp Tyr  
 20 25 30  
 Val His Asn Leu Lys Lys Phe Leu Ser Glu Gly Gln Ile Val Lys Ala  
 35 40 45  
 30 Lys Ile Xaa Ser Ile Asp Asp Glu Gly Lys Leu Asn Leu Ser Leu Xaa  
 50 55 60  
 Asp Asn Asp Tyr Phe Lys Asn Tyr Glu Arg Lys Lys Glu Lys Gln Ser  
 65 70 75 80  
 35 Val Leu Asp Glu Ile Arg Xaa Thr Glu Lys Tyr Gly Xaa Gln Thr Leu  
 85 90 95  
 Lys Arg Thr Leu Thr Asn Leu Gly  
 100

(2) INFORMATION FOR SEQ ID NO:266:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 97 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:266:

EP 0 841 394 A2

Val Thr Xaa Gly Val Xaa Glu Leu Ile Ile Xaa Ser Gly Glu Glu Glu  
 1 5 10 15  
 Ser Ser Ser Gln Leu Cys His Pro His Ser Val Lys Asp Val Thr Gly  
 5 20 25 30  
 Ala Gly Asp Ser Phe Cys Ala Ala Val Val Tyr Ser Trp Xaa Asn Gly  
 35 40 45  
 Met Ser Thr Val Asp Ile Leu Ile Ala Gly Met Val Asn Ala Xaa Lys  
 10 50 55 60  
 Thr Ile Glu Thr Lys Tyr Thr Val Arg His Asn Leu Asp Gln Gln Gln  
 65 70 75 80  
 Leu Tyr His Asp Met Glu Asp Tyr Lys Asn Gly Lys Phe Thr Gln Val  
 15 85 90 95  
 Tyr

(2) INFORMATION FOR SEQ ID NO:267:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 56 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:267:

Val Asn Lys Asn Xaa Thr Thr Met Ser His Ile Thr Xaa Xaa Gln Ser  
 1 5 10 15  
 Ala Ile Leu Ala Ser Lys Val Asn Ala Pro Ser Val Tyr Asn Xaa Asn  
 20 25 30  
 Asn Met Ser Glu Asn Phe Thr Gln Arg Val Lys His Xaa Leu Arg Lys  
 35 40 45  
 Asn Glu Ala Thr Thr Xaa Tyr Gln  
 50 55

(2) INFORMATION FOR SEQ ID NO:268:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 54 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

# EP 0 841 394 A2

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:268:

```

Val Ser Ser Arg Asp Ile Gly Glu His Val Met Asn Leu Leu Met His
 1           5           10           15
Val Asp Gln Val Ser Tyr Val Arg Phe Ala Ser Val Tyr Lys Glu Phe
10          20          25          30
Lys Asp Val Asp Gln Leu Leu Ala Ser Met Gln Gly Ile Leu Ser Glu
          35          40          45
Asn Lys Arg Ser Asp Ala
15          50

```

(2) INFORMATION FOR SEQ ID NO:269:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 104 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:269:

```

Val Ala Xaa Val Xaa Ile Ile Xaa Phe Met Met Leu Xaa Pro Xaa Gly
 1           5           10           15
Gly Phe Ile Ala Leu Glu Gln Ile Gly Glu Arg Met Asn Ile Ala Ala
35          20          25          30
Ile Glu Val Asp Lys Ser Tyr Arg Phe Asn Gly Ile Gly Ser Ser Ala
          35          40          45
Leu Arg Gln Leu Pro Thr Tyr Leu Arg Lys Asn Tyr Asp Asn Leu Asn
40          50          55          60
Val Ile Thr Met Ile Leu Phe Gly Glu Asn Asn Asp Phe Lys Pro Leu
          65          70          75          80
Cys Leu Asn Ser Asn Phe Val Glu Ile Glu Gln Thr Asp Asp Tyr Val
45          85          90          95
Val Phe Glu Lys Tyr Leu Asn Tyr
          100

```

(2) INFORMATION FOR SEQ ID NO:270:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 90 amino acids

EP 0 841 394 A2

(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:270:

```

Val Glu Ile Lys Val His Ala Leu His Ser Gly Ile Thr Leu Ile Lys
 1             5             10             15
Asp Lys Gly Lys Ile Ile Asp Ile His Leu Ser Val Lys Ala Thr Glu
15          20          25          30
Asn Ile Asp Gly Glu Val Leu Phe Lys Ala Thr Gln Pro Leu Gly Arg
          35          40          45
Thr Met Lys Val Gly Val Gln Asn Asn Ala Met Thr Ile Thr Leu Thr
20          50          55          60
Lys Gln Asn Gln Trp Leu Asp Ser Leu Lys Phe Leu Val Lys Cys Ile
65          70          75          80
Glu Glu Ser Met Arg Ile Ser Asp Glu Ala
25          85          90

```

(2) INFORMATION FOR SEQ ID NO:271:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 114 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:271:

```

Val Glu Tyr Phe Val Ser Tyr Tyr Asp Tyr Tyr Gln Pro Glu Ala Tyr
 1             5             10             15
Val Pro Ser Thr Asp Thr Phe Ile Glu Lys Asp Ala Ser Ile Asn Asp
45          20          25          30
Glu Ile Asp Xaa Xaa Arg His Ser Ala Thr Ser Ala Leu Phe Glu Arg
          35          40          45
Asp Asp Val Ile Ile Ile Ala Ser Val Ser Cys Ile Tyr Gly Leu Gly
50          50          55          60
Asn Pro Glu Glu Tyr Lys Asp Leu Val Val Ser Val Arg Val Gly Met
65          70          75          80

```

EP 0 841 394 A2

... Glu Met Asp Arg Ser Glu Leu Leu Arg Lys Leu Val Asp Val Gln Tyr ...  
 85 90 95  
 5 Thr Arg Asn Asp Ile Asp Phe Pro Thr Arg Asn Arg Phe Glu Cys Val  
 100 105 110  
 Val Met

10

(2) INFORMATION FOR SEQ ID NO:272:

(i) SEQUENCE CHARACTERISTICS:

15

- (A) LENGTH: 52 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:272:

25

Val Gly His Ile Leu Ile Glu Ile Ser Xaa Asp Gly Pro Gly Phe Met  
 1 5 10 15  
 Glu Asp Xaa Pro Tyr Glu Thr Leu Gly Xaa Gly Leu Ser Leu Pro Pro  
 20 25 30  
 Phe Leu Glu Asn Lys Arg Glu Tyr Ile Glu Ser Glu Ile Arg Pro Phe  
 35 40 45  
 Asn Thr Lys Arg  
 50

35

(2) INFORMATION FOR SEQ ID NO:273:

(i) SEQUENCE CHARACTERISTICS:

40

- (A) LENGTH: 46 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

45

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:273:

50

Val Ser Ile Thr Leu Xaa Pro Ala Leu Ala Ala Thr Leu Phe Lys Lys  
 1 5 10 15  
 Gly Val Lys Arg Arg Asn Lys Gln His Gln Glu Gly Leu Gly Val Val  
 20 25 30

55

EP 0 841 394 A2

... Ser Thr Thr Tyr Lys Xaa Val Xaa His Trp Val Thr Lys Ser ...  
 35 40 45

(2) INFORMATION FOR SEQ ID NO:274:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:274:

Val Val Gln Glu Thr Leu Asn Met Ser Lys Ala Arg Gly Tyr Glu Val  
 1 5 10 15  
 Gly Gly Thr Val Arg Ile Val Ile Asn Asn Gln Val Gly Phe Thr Thr  
 20 25 30  
 Ser Asn Pro Leu Asp Ala Arg Ser Tyr Ala Val Leu Tyr  
 35 40 45

(2) INFORMATION FOR SEQ ID NO:275:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 88 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:275:

Val Glu Gly Asn Phe Gly Leu Ala Ile Asn Asn His Gln Asn Ile Val  
 1 5 10 15  
 Gly Thr Ile Gly Leu Ile Arg Leu Asp Asn Asn Met Ser Ala Leu Lys  
 20 25 30  
 Lys Met Phe Val Asp Lys Gly Tyr Arg Asn Leu Lys Ile Gly Lys Lys  
 35 40 45  
 Leu Leu Asp Lys Val Ile Met Thr Cys Lys Glu Gln Asn Ile Asp Xaa  
 50 55 60  
 Xaa Tyr Leu Gly Thr Xaa Asp Lys Xaa Xaa Ser Ala Gln Tyr Xaa Tyr  
 65 70 75 80

EP 0 841 394 A2

Xaa Asn Xaa Gly Phe Arg Xaa Asn  
85

5

(2) INFORMATION FOR SEQ ID NO:276:

(i) SEQUENCE CHARACTERISTICS:

10

(A) LENGTH: 96 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:276:

20

Val	Ile	Gly	Leu	Lys	Ser	Lys	Thr	Ile	Ile	Ser	Asn	Glu	Trp	Gly	Ala
1				5				10						15	
Ile	Met	Ile	Gln	Ser	Met	Trp	Phe	Asn	Leu	His	Val	Gln	Asp	Leu	Glu
			20					25					30		
Lys	Ser	Ala	Gln	Phe	Tyr	Lys	Ala	Leu	Gly	Phe	Lys	Ile	Asn	Arg	Asn
		35					40					45			
Pro	Gln	Met	Leu	Asp	Lys	Met	Val	Gly	Ile	Gln	Ile	Gly	Gln	Thr	Thr
	50					55					60				
Ala	Ile	Leu	Ile	Glu	Asn	Lys	His	Phe	Gln	Asn	Val	Ser	Gln	Gln	Ser
65					70					75				80	
Leu	Asn	Thr	Glu	Pro	Asn	Glu	Val	Met	Ile	Ser	Leu	Gly	Val	Thr	Gln
			85					90						95	

35

(2) INFORMATION FOR SEQ ID NO:277:

(i) SEQUENCE CHARACTERISTICS:

40

(A) LENGTH: 85 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

45

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:277:

50

Val	Ser	Leu	Met	Thr	Thr	Phe	Ser	Glu	Lys	Glu	Lys	Ile	Gln	Leu	Leu
1				5				10					15		
Ala	Asp	Ile	Val	Glu	Leu	Gln	Thr	Glu	Asn	Asn	Asn	Glu	Ile	Asp	Val
			20					25					30		

55

EP 0 841 394 A2

... Cys Asn Tyr Leu Lys Asp Leu Phe Asp Lys Tyr Asp Ile Lys Ser Glu ...  
 35 40 45  
 5 Ile Leu Lys Val Asn Glu His Arg Ala Asn Ile Val Ala Glu Ile Gly  
 50 55 60  
 Asn Xaa Ser Pro Ile Leu Ala Leu Ser Gly His Met Asp Val Val Gly  
 65 70 75 80  
 10 Cys Arg Lys Ser Arg  
 85

(2) INFORMATION FOR SEQ ID NO:278:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 66 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:278:

Val Glu Xaa Met Asn Lys His Tyr Xaa Ile Val Ile Ile Gly Gly Gly  
 1 5 10 15  
 30 Thr Ala Gly Val Thr Val Ala Ser Arg Leu Leu Arg Lys Asn Xaa Asn  
 20 25 30  
 Leu Lys Glu Lys Ile Ala Ile Ile Asp Pro Ala Asp His His Tyr Tyr  
 35 35 40 45  
 Gln Pro Leu Trp Thr Leu Val Gly Ala Gly Val Ser Ser Leu Lys Ser  
 50 55 60  
 Xaa Ser  
 65  
 40

(2) INFORMATION FOR SEQ ID NO:279:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 61 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:279:



EP 0 841 394 A2

Val Leu Arg Asn Lys Ile Lys Arg Ala Ile Arg Glu Asn Phe Lys Val  
 1 5 10 15  
 5 His Lys Ser Xaa Ile Leu Xaa Lys Asp Ile Ile Xaa Ile Xaa Arg Gln  
 20 25 30  
 Ala Ala Lys Asp Met Thr Thr Leu Gln Ile Gln Xaa Ser Leu Glu His  
 35 40 45  
 10 Xaa Leu Lys Ile Gly Lys Val Phe Asn Lys Lys Ile Lys  
 50 55 60

(2) INFORMATION FOR SEQ ID NO:280:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 71 amino acids  
 (B) TYPE: amino acid  
 20 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:280:

Val Asp Gly Val Ser Leu Ala Val Phe Asp Lys His Asp Asn Ser Phe  
 1 5 10 15  
 30 Asp Ile His Leu Ile Pro Glu Thr Arg Arg Ser Thr Ile Leu Ser Ser  
 20 25 30  
 Thr Lys Leu Gly Asp Lys Val His Leu Glu Thr Asp Val Leu Phe Lys  
 35 35 40 45  
 Tyr Val Glu Asn Ile Leu Asn Lys Asp Lys Asp Gln Leu Ser Val Asp  
 50 55 60  
 Lys Leu Arg Ala Phe Xaa Phe  
 65 70  
 40

(2) INFORMATION FOR SEQ ID NO:281:

45 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 126 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 50 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:281:

EP 0 841 394 A2

Val Xaa Gln Ser Leu Tyr Glu Phe Leu Glu Glu Asn Ile Asn Tyr Leu  
 1 5 10 15  
 5 Lys Glu Asn Gly Xaa Tyr Asn Glu Ile Asp Thr Ile Glu Gly Ala Asn  
 20 25 30  
 Gly Pro Glu Ile Lys Ile Asn Gly Lys Ser Tyr Ile Asn Leu Ser Ser  
 35 40 45  
 10 Xaa Asn Tyr Leu Gly Leu Ala Thr Asn Glu Asp Xaa Arg Ser Ala Ala  
 50 55 60  
 Lys Ala Ala Ile Asp Thr His Gly Val Gly Ala Gly Ala Xaa Arg Thr  
 65 70 75 80  
 15 Xaa Asn Gly Thr Leu Asp Leu His Asp Glu Leu Glu Glu Thr Leu Ala  
 85 90 95  
 Lys Phe Xaa Gly Thr Glu Ala Ala Ile Ala Tyr Gln Xaa Arg Ile Xaa  
 100 105 110  
 20 Leu Xaa Ile Trp Leu Leu Xaa Gln Leu Val Xaa Glu Xaa Lys  
 115 120 125

(2) INFORMATION FOR SEQ ID NO:282:

25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 51 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 30 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:282:

Val Leu Gly Gly Asn Val Thr Glu Thr His Thr Phe Lys Leu Pro Glu  
 1 5 10 15  
 40 Asp Ala Gly Glu Arg Gln Met Phe Ile Ile Asp Lys Lys Arg Gln Thr  
 20 25 30  
 Pro Lys Lys Tyr Pro Arg Lys Pro Gly Thr Pro Asn Lys Thr Pro Leu  
 35 40 45  
 45 Leu Glu Lys  
 50

(2) INFORMATION FOR SEQ ID NO:283:

50 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 32 amino acids  
 (B) TYPE: amino acid  
 55 (C) STRANDEDNESS: single

EP 0 841 394 A2

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:283:

Val Lys Glu Asn Asp Val Lys Lys Phe Lys Tyr Lys Asn Phe Xaa Asp  
 1 5 10 15  
 Glu Ile Glu Ile Asp Phe Thr Asp Ser Asn His Leu Ala Ala His Arg  
 20 25 30

(2) INFORMATION FOR SEQ ID NO:284:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 76 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:284:

Val Thr Pro Asp Ala Asn Gly Lys Val Ala Phe Asp Gly Leu Glu Leu  
 1 5 10 15  
 Thr Phe Thr Gly Thr Pro Ala Val Asn Asp Ser Phe Thr Leu Lys Pro  
 20 25 30  
 Val Ser Asp Ala Ile Val Asn Met Asp Val Leu Ile Thr Asp Glu Ala  
 35 40 45  
 Lys Ile Ala Met Ala Ser Glu Glu Asp Ala Gly Asp Ser Asp Asn Arg  
 50 55 60  
 Asn Gly Gln Ala Leu Leu Asp Leu Ala Lys Gln Gln  
 65 70 75

(2) INFORMATION FOR SEQ ID NO:285:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 73 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:285:

```

5   Val Leu Leu Pro Leu Val Phe Ile Ser Val Leu Ile Gly Ile Phe Asn
    1             5             10             15
    Tyr Ile Lys Val Leu Pro Phe Ile Ile Lys Tyr Val Gly Ile Ala Ile
        20             25             30
10  Asn Lys Ile Thr Arg Met Gly Arg Leu Glu Ser Tyr Phe Ala Ile Ser
    35             40             45
    Thr Ala Met Phe Gly Gln Pro Glu Val Tyr Leu Thr Ile Lys Asp Ile
    50             55             60
15  Ile Pro Arg Leu Ser Arg Ala Lys Leu
    65             70

```

(2) INFORMATION FOR SEQ ID NO:286:

```

20  (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 54 amino acids
    (B) TYPE: amino acid
25  (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

```

(ii) MOLECULE TYPE: Protein

```

30  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:286:

```

```

    Val Ile Ile Ser Asn Asn Val Gly Asp Met Ser Ile Gly Phe Ser Ser
35   1             5             10             15
    Asp Asp Asn Glu Leu Thr Met His Phe Arg Asn Asn Glu Lys Val Asn
        20             25             30
    Ile Lys Lys Gly Lys Lys Val Val Leu Ala Ala Gln Ile Leu Asp Glu
40   35             40             45
    Leu Glu Thr Arg Trp Gln
    50

```

45 (2) INFORMATION FOR SEQ ID NO:287:

```

    (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 88 amino acids
50  (B) TYPE: amino acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

```

```

55  (ii) MOLECULE TYPE: Protein

```

EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:287:

```

5   Val Thr Tyr Xaa His Met Ser His Arg His Thr Leu Ile Ser His Ser
    1           5           10           15
    Phe Lys Phe Xaa Arg Leu Pro Gln Asp Ile Asn Phe Phe Ser Trp Xaa
      20           25           30
10  Gln His Val Lys Asp Ser Asp Lys Thr Asp Glu Leu Thr Tyr Ser Glu
    35           40           45
    Ser Leu Xaa Pro Glu Gly His Pro Thr His Pro Leu Thr Pro Thr Lys
    50           55           60
15  Leu Pro Xaa Thr Met Glu Glu Xaa Arg Ala Tyr Ala Pro Glu Phe Asp
    65           70           75           80
    Xaa Arg Asn Pro Phe Ala Thr Leu
      85

```

(2) INFORMATION FOR SEQ ID NO:288:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 70 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:288:

```

35  Val Ile Ser Leu Leu Lys Asp Ala Lys Leu Pro Ser Ser Val Thr Arg
    1           5           10           15
    Gly Asp Tyr Leu Ala Ile Leu Ser Xaa Gly Ala Tyr His Tyr Ser Met
      20           25           30
40  Xaa Ser Asn Tyr Asn His Met His Lys Pro Ser Val Phe Phe Leu Lys
    35           40           45
    Asp Val Thr Ala Arg Val Val Ile Lys Arg Gln Ser Leu Arg Gln Leu
    50           55           60
45  Ile Ile Asn Xaa Thr Gln
    65           70

```

(2) INFORMATION FOR SEQ ID NO:289:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 67 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single

EP 0 841 394 A2

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:289:

```

10 Val Gln Phe Val Asn Lys Leu Phe Asn Val Ser Ser Ala Ile Ile Leu
    1           5           10           15
    Leu Glu Tyr Asp Gly Val Val His Ile Gly Tyr Asp Asn Asn Phe Glu
        20           25           30
15 Phe Lys Thr Glu Gln Phe Lys Met Ser Lys Ser Arg Asn Leu Leu Lys
    35           40           45
    Asn Arg Ser Gln Asn Xaa Val Leu Ile Arg Leu Leu Asn Trp Leu Arg
    50           55           60
20 Thr Thr Xaa
    65

```

(2) INFORMATION FOR SEQ ID NO:290:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 41 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:290:

```

40 Val Gln Leu Gly Leu Thr Asn Ser Val Val Ala Ile Ile Ile Ile Xaa
    1           5           10           15
    Arg Glu Xaa Ala Val Thr Gly Leu Arg Leu Leu Gln Ile Glu Gln Gly
        20           25           30
45 Xaa Gly Lys Cys Ser Trp Xaa Ile Arg
    35           40

```

(2) INFORMATION FOR SEQ ID NO:291:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 66 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

EP 0 841 394 A2

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:291:

Val Asp Val Asp Asn Met Ser Asp Tyr Lys Leu Lys Ile Ile Xaa Leu  
 1 5 10 15  
 Thr Xaa Ser Asp Ile Thr Gly Tyr Gln Ile Pro Asn Gln Xaa Gly Val  
 20 25 30  
 Ala Gln Tyr Val Ile Ser Gln Leu Ser Gln Gly Lys Arg Glu Val Asp  
 35 40 45  
 Asn Leu Thr Leu Asn Xaa Xaa Glu Xaa Leu Tyr Ser Tyr Xaa Arg Gln  
 50 55 60  
 Val Leu  
 65

(2) INFORMATION FOR SEQ ID NO:292:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:292:

Val Lys Gln Thr Xaa Val Glu Glu Ile Xaa Ala Ser Ile Asn Xaa Ala  
 1 5 10 15  
 Gly Leu Xaa Xaa Glu Ile Pro Asp Phe Lys Xaa Glu Val Xaa Xaa Xaa  
 20 25 30  
 Xaa Lys Lys Trp Arg Leu Cys Tyr  
 35 40

(2) INFORMATION FOR SEQ ID NO:293:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 106 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:293:

```

5   Val Ser Thr Asp Phe Ile Leu Cys Lys Leu Gln Ala Phe His Ile Ile
    1             5             10             15
    Lys Phe Glu Lys Arg Tyr Ile Glu Val Glu Lys Asn Glu Tyr Thr Ala
      20             25             30
10   Lys Tyr Asn Glu Tyr Ser Gln Leu Leu Asp Ala Thr Tyr Ser Gln Ala
      35             40             45
    Val Ala Tyr Leu Leu Asn Lys Tyr Gly Ala Val Thr Asp Asp Tyr Tyr
      50             55             60
15   Lys Glu Lys Ser Tyr Thr Arg Phe Leu Asn Gly Glu Ile Lys Ser Ile
      65             70             75             80
    Ser Lys Gly Lys Tyr Thr Arg Ala Ser Glu Gly Leu Tyr Cys His His
      85             90             95
20   Ile Ser Glu Asp Lys Phe Gln Asn Leu Ser
      100             105

```

(2) INFORMATION FOR SEQ ID NO:294:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 86 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:294:

```

40   Val Ile Leu Ser Ser Ile Xaa Phe Tyr Met Gln Leu Leu Gln Lys Asn
    1             5             10             15
    Pro Met Ala Ile Gly Val Trp Ala Gly Leu Cys Gln Met Ile Lys Pro
      20             25             30
    Thr Val Ile Asp Trp Asp Ile Ser Glu Tyr Thr Pro Lys Pro Ala Trp
      35             40             45
45   Met Gln Ala Thr Lys Ala Arg Ala Tyr Val Pro Arg Ile Leu Val Glu
      50             55             60
    Lys Leu Leu Ile Xaa Ile Asp Asp Met Leu Glu Gly Ile Glu Ile Tyr
      65             70             75             80
50   Asp Xaa Xaa Glu Ser Arg
      85

```

(2) INFORMATION FOR SEQ ID NO:295:



EP 0 841 394 A2

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 82 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:295:

```

Val Xaa Thr Tyr Val Asn Asp Tyr Leu Asn Glu Ala His Ile Thr Asn
 1             5             10             15
Lys Trp Ser Glu Met Met Leu Trp Ser Ser Gln Gln Arg Ser Gln Tyr
          20             25             30
Thr Val Gln Leu Ile Asn Lys Ile Glu Thr Asp Asp Ser Tyr Ile His
      35             40             45
Ile Ser Lys Gly Glu Leu Gly Ile Ser Leu Met Pro His Ile Gln Leu
      50             55             60
Lys Lys Ala Met Ser Asn Thr Ala Ser His Ile Glu Asp Leu Ser Cys
25 65             70             75             80
Gly Leu

```

(2) INFORMATION FOR SEQ ID NO:296:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 59 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:296:

```

Val Cys Lys Lys Tyr Val Leu Gly Gln Phe Lys Pro Gly Met Thr Ala
 1             5             10             15
Pro Gln Cys Ala Gly Ile Ile His Thr Asp Phe Xaa Arg Gly Phe Ile
      20             25             30
Arg Ala Gly Ser Asn Lys Leu Leu Met Thr Met Tyr Asn Met Ala Ala
      35             40             45
Lys Val Ala Leu Lys Lys Arg Ala Asp Thr Asp
55 50             55

```

EP 0 841 394 A2

(2) INFORMATION FOR SEQ ID NO:297:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 56 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:297:

```

Val Leu Thr Leu Ala Glu Met Lys Arg Thr Ile His Asp Ile Leu Asp
 1             5             10             15
Phe Arg Asp Glu Asp Ile Trp Xaa Cys Tyr Leu Gly Thr Leu Ala Val
20             20             25             30
Ser Pro Xaa Leu Xaa Asp Asp Xaa Gly Xaa Xaa Leu Leu Ser Arg Xaa
35             40             45
Xaa Asn Ala Tyr Asn Xaa Tyr Xaa
25             50             55

```

(2) INFORMATION FOR SEQ ID NO:298:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 98 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:298:

```

Val Phe Asn Tyr Gly Gln Ser Val Phe Glu Gly Leu Lys Ala Tyr Lys
45  1             5             10             15
Arg Asp Gly Glu Gly Cys Thr Xaa Pro Xaa Pro Glu Glu Asn Phe Lys
20             25             30
Arg Leu Asn Asn Ser Leu Ala Arg Leu Glu Met Pro Gln Val Asp Asp
50  35             40             45
Ala Glu Leu Leu Glu Gly Leu Lys Xaa Leu Val Asp Ile Glu Arg Asp
55  50             55             60
Xaa Xaa Pro Glu Gly Glu Gly Gln Ser Leu Tyr Xaa Xaa Pro Phe Gly
65             70             75             80

```

EP 0 841 394 A2

Xaa Ala Thr Glu Gly Ala Xaa Gly Val Gly Ala Ser His Gln Tyr Xaa  
 85 90 95  
 5 Ile Tyr

(2) INFORMATION FOR SEQ ID NO:299:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:299:

Val Arg Phe Ile Leu Glu Asp Pro Pro Arg Asp Leu Asp Ile Tyr Ile  
 1 5 10 15  
 Asp Ile Ala Glu Arg Ala Gly Gln His Pro Xaa Val Glu Arg Ile Xaa  
 20 25 30  
 Asp Gly Xaa Gln Arg Arg Ile Asp Ile Xaa Thr Val Glu Ser Cys  
 35 40 45

(2) INFORMATION FOR SEQ ID NO:300:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 123 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:300:

Val Val Lys Thr Arg Ile Leu Lys Val Asp Asn Pro Glu Ile Thr Ile  
 1 5 10 15  
 Pro Leu Lys Pro Thr Gly Ser His Phe Gln Gln Cys Val Trp Asn Glu  
 20 25 30  
 Leu Arg Gln Val Pro Tyr Gly Thr Leu Thr Thr Tyr Gly Ala Ile Ala  
 35 40 45  
 Thr Lys Val Gly Lys Val Leu Asp Lys Pro Gln Met Ser Ala Gln Ala  
 50 55 60

EP 0 841 394 A2

... Val Gly Gly Ala Val Gly Ser Asn Pro Leu Ser Ile Ile Val Pro Cys ...  
 65 70 75 80  
 5 His Arg Val Val Gly Lys Thr Gly Ser Leu Thr Gly Phe Gly Gly Thr  
 85 90 95  
 Ile Asn Asn Lys Ile Lys Leu Leu Glu Leu Glu Asn Ile Asp Met Ser  
 100 105 110  
 10 Lys Leu Tyr Ile Pro Lys His Ser Thr Lys Pro  
 115 120

(2) INFORMATION FOR SEQ ID NO:301:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 62 amino acids  
 (B) TYPE: amino acid  
 20 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:301:

Val Cys Leu Val Glu Cys Met Met Ala Leu Asn Thr Ala Gly Xaa Thr  
 30 1 5 10 15  
 Gln Arg Thr Val Glu Thr Leu Ala Glu Tyr Ser Gly Val Pro Val Trp  
 20 25 30  
 Asn Gly Leu Thr Asp Glu Asp His Pro Thr Gln Xaa Leu Ala Asp Phe  
 35 35 40 45  
 Leu Thr Ala Lys Glu Val Phe Lys Lys Arg Leu Cys Arg Tyr  
 50 55 60

(2) INFORMATION FOR SEQ ID NO:302:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 88 amino acids  
 45 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:302:

55 Val Asn Leu Pro Asp Phe Pro Pro Lys Ile Gly Val Asn Lys Ser Thr  
 1 5 10 15

EP 0 841 394 A2

Leu Ser Arg Tyr Xaa Xaa Gly Ser Arg Xaa Ile Pro Met Glu Asp Ile  
 20 25 30  
 5 Ala Glu Ile Ala Asn Ala Leu Lys Val Thr Pro Glu Tyr Leu Leu Leu  
 35 40 45  
 Xaa Asn Arg Gln Pro Glu Xaa Glu Val Gln His Arg Ala Ala His Leu  
 50 55 60  
 10 Glu Gly Glu Leu Thr Asp Asp Glu Cys Gln Arg Val Leu Asp Tyr Ala  
 65 70 75 80  
 Asp Tyr Ile Arg Ser Lys Arg Lys  
 85

(2) INFORMATION FOR SEQ ID NO:303:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 121 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:303:

Val Lys Pro Xaa Cys Ala Lys Lys Asp Arg Tyr Pro Ile Asp Pro Met  
 1 5 10 15  
 Arg Arg Leu Asn Gln Glu Ile Ile Asp Tyr Ile Asp Thr Leu Asn Tyr  
 20 25 30  
 35 Ile Asp Gln Tyr Ser Ser Ala Gly Ser Phe Arg Arg Phe Lys Glu Met  
 35 40 45  
 Ser Lys Asp Leu Asp Phe Ile Ile Ser Thr Asp Asn Pro Gln Ala Val  
 40 50 55 60  
 Gln Gln Gln Leu Leu Asn Ile Pro Asn Lys Val Lys Glu Val Ala Val  
 65 70 75 80  
 Gly Asn Thr Gln Val Ser Leu Glu Leu Ala Tyr Asp Asp Glu Thr Ile  
 45 85 90 95  
 Gly Val Asp Phe Arg Leu Ile Glu Pro Arg Ala Phe Tyr His Thr Leu  
 100 105 110  
 Gln His Phe Thr Gly Ser Asn Arg Thr  
 50 115 120

(2) INFORMATION FOR SEQ ID NO:304:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 116 amino acids

EP 0 841 394 A2

(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:304:

```

Val Phe Asp Ile Tyr Ser Glu Xaa Asp Leu Ile Glu Ile Thr Asp Tyr
 1             5             10             15
Pro Xaa Ser Ile Ser Phe Asn Phe Ala Ser Tyr Met Ser Ala Ser Lys
15            20            25            30
Phe Phe Lys Asp Tyr Ala Leu Xaa Thr Asn Asp Lys Ser Gln Tyr Leu
35            40            45
Glu Asp Tyr Asn Gln His Val Ala Ile Val Ala Leu Tyr Leu Ala Xaa
20            50            55            60
Gly Asn Lys Ala Gln Ala Lys Gln Phe Ile Ser Ala Met Val Glu Gln
65            70            75            80
Arg Tyr Xaa Pro Ala Thr Pro Thr Phe Leu Asn Ala Xaa Arg Ala Arg
25            85            90            95
Arg Gly Glu Leu Val Ser Cys Phe Xaa Xaa Lys Lys Trp Val Xaa Xaa
100           105           110
Leu Asn Ser Xaa
30           115

```

(2) INFORMATION FOR SEQ ID NO:305:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 266 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:305:

```

Val Met Glu Val Thr Cys Leu Lys Gln Ile Thr Trp His Asp Leu Gln
50  1             5             10             15
His Ile Ile Lys Asp Gly Asp Val Ile Gly Leu Pro Ala Leu Ala Val
20            25            30
Ala Asn Leu Pro Ala Glu Val Leu Arg Ala Val Leu Ala Gln His Asp
55            35            40            45

```

EP 0 841 394 A2

```

... Thr Tyr His Thr Pro Lys Asp Leu Thr Phe Ile Leu Ala Asn Asp Ile ...
    50          55          60
5   His Ser Leu Gly Ala Ala Pro Asp Leu Asp Asp Phe Ile Glu Arg Arg
    65          70          75          80
    Met Ile Lys Arg Val Ile Met Ser Ile Leu Thr Ala Ser Ser Lys Thr
          85          90          95
10  Ala Gln Ala Met Lys Asn Asn Asp Ile Glu Ala Tyr Phe Leu Pro Gln
          100         105         110
    Gly Ile Ile Ala Thr His Tyr Arg Gln Ser Asn Gln Leu Leu Pro Gly
          115         120         125
15  Val Ile Thr Lys Ile Gly Leu Asn Thr Ala Val Asp Pro Arg Tyr Gly
    130         135         140
    Gly Gly Lys Val Asn Thr Arg Thr Thr Asp Asp Leu Val Ser Leu Val
    145         150         155         160
20  Thr Ile Asn Asp Glu Thr Tyr Leu His Tyr Thr Phe Pro Ser Val Asp
          165         170         175
    Val Ala Leu Leu Arg Gly Thr Tyr Ala Asp Gln Gln Gly Asn Ile Tyr
          180         185         190
25  Leu Xaa Gln Glu Ala Tyr Leu Ser Glu Cys Tyr His Val Ala Leu Asn
    195         200         205
    Ala Lys Ala Asn His Gly Lys Val Ile Xaa Gln Val Lys Ala Leu Val
    210         215         220
30  Asp Asp Tyr His Leu Lys Pro Asn Glu Val Val Ile Pro Gly Asn Leu
    225         230         235         240
    Val Asp Tyr Val Phe Val Thr Glu Asp Glu Lys Asn His Arg Gln Val
          245         250         255
35  Ile Gln Ser His Tyr Leu Pro Ala Leu Ser
          260         265

```

(2) INFORMATION FOR SEQ ID NO:306:

```

40  (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 68 amino acids
      (B) TYPE: amino acid
45  (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

```

(ii) MOLECULE TYPE: Protein

```

50  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:306:

```

```

55  Val Val Lys Ile Val Xaa His Gln Lys Asp Ser Asn Xaa Glu Lys Asp
    1          5          10          15

```

EP 0 841 394 A2

Asp Lys Ala Arg Thr Xaa Gln Gln Asp Xaa Lys Xaa Xaa Asp Ser Ser  
 20 25 30  
 Xaa Asp Lys Lys Asp Asn Xaa Asp Asp Ser Xaa Asp Val Xaa Lys Asp  
 5 35 40 45  
 Asn Lys Asp Asn Ser Ala Asn Asp Asn Gln Gln Gln Ser Asn Ser Lys  
 50 55 60  
 Cys Asn Lys Gln  
 10 65

(2) INFORMATION FOR SEQ ID NO:307:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 171 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:307:

Val Arg Arg Phe Pro Asp Ile Glu Arg Arg Val Gln Thr His Asp Phe  
 1 5 10 15  
 Thr Trp Gln Ala Glu Ile Met Ser Lys Pro Val Lys Asn Met Thr Val  
 20 25 30  
 Ala Ile Ile Gly Thr Gly Arg Ile Gly Ala Ala Thr Ala Lys Ile Tyr  
 35 40 45  
 Ala Gly Phe Gly Ala Thr Ile Thr Ala Tyr Asp Ala Tyr Pro Asn Lys  
 50 55 60  
 Asp Leu Asp Phe Leu Thr Tyr Lys Asp Ser Val Lys Glu Ala Ile Lys  
 65 70 75 80  
 Asp Ala Asp Ile Ile Ser Leu His Val Pro Ala Asn Lys Glu Ser Tyr  
 85 90 95  
 His Leu Phe Asp Lys Ala Met Phe Asp His Val Lys Lys Gly Ala Ile  
 100 105 110  
 Leu Val Asn Ala Ala Arg Gly Ala Val Ile Asn Thr Pro Asp Leu Xaa  
 115 120 125  
 Ala Ala Val Asn Asp Gly Thr Leu Leu Gly Ala Ala Ile Asp Thr Tyr  
 130 135 140  
 Glu Asn Glu Ala Ala Tyr Phe Thr Asn Asp Trp Thr Asn Lys Asp Ile  
 145 150 155 160  
 Asp Asp Lys Thr Leu Leu Glu Leu Ile Gly Thr  
 165 170



EP 0 841 394 A2

(2) INFORMATION FOR SEQ ID NO:308:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 52 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:308:

```

Val Val Val Leu Met Cys Cys Leu Lys Tyr Glu Asn Asp Tyr Tyr Glu
 1             5             10             15
Glu Val Arg Ala Gln Leu Pro Asp Ile Gly Glu Ala Ile Glu Thr Pro
20             25             30
Asp Gly Asn Arg Glu Ser Ser Cys Phe Lys Tyr Ile Arg His Phe Tyr
35             40             45
Ala Gly Glu Ala
25             50

```

(2) INFORMATION FOR SEQ ID NO:309:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 114 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:309:

```

Val Asn Thr Xaa Leu Leu Asn Xaa Arg Val Glu Pro Ala Ile Val Lys
45  1             5             10             15
Ile Pro Val Ser Val Asp Glu Asn Ile Asp Asn Val Glu Lys Lys Leu
20             25             30
Asn Thr Leu Phe Thr Ile Leu Cys Val Val Asn Xaa Xaa Leu Phe Val
50  35             40             45
Ser Asp Pro Val Val Ile Gly Ile Asp Ala Phe Glu Asp Thr Arg Val
50             55             60
Ile Leu Xaa Val Ser Ala Glu Thr Ile Pro Gly Xaa Gly Phe Ser Gly
55  65             70             75             80

```

EP 0 841 394 A2

Ala Arg Ile Ile Arg Lys Glu Val His Lys Met Phe Leu Gln Glu Gly  
 85 90 95  
 5 Ile Lys Thr Pro Xaa Pro Ile Met Thr Pro Phe Asn His Ser Glu Xaa  
 100 105 110  
 Gly Val

(2) INFORMATION FOR SEQ ID NO:310:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:310:

Val Thr Gln Leu Leu Lys Asn Ala Xaa Lys Lys Xaa Asn Asp Leu Ile  
 1 5 10 15  
 Lys Glu Glu Ala Glu Ile Val Lys Asn Asp Asp Leu Pro Xaa Xaa Leu  
 20 25 30  
 Ile Ile Xaa Gly Xaa Gln Tyr Asp Xaa Arg Trp Arg  
 35 40

(2) INFORMATION FOR SEQ ID NO:311:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 119 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:311:

Val Arg Phe His Asp Gly Xaa Thr Phe Asp Ala Asp Ala Val Lys Lys  
 1 5 10 15  
 Asn Ile Asp Ala Xaa Gln Gln Asn Lys Lys Leu His Ser Trp Leu Lys  
 20 25 30  
 Ile Ser Thr Leu Ile Asp Asn Val Lys Val Lys Asp Lys Tyr Thr Val  
 35 40 45

EP 0 841 394 A2

[illegible]

(2) INFORMATION FOR SEO ID NO:312:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 307 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:312:

Val	Pro	Ile	Asp	Tyr	Ile	Xaa	Gln	Thr	Gly	Asp	Ile	Val	Glu	Ile	Arg
1				5					10					15	
Thr	Ser	Lys	His	Ser	Tyr	Gly	Pro	Ser	Arg	Asp	Trp	Leu	Xaa	Ile	Val
			20					25					30		
Lys	Ser	Ser	Ser	Ala	Lys	Gly	Lys	Ile	Lys	Ser	Phe	Phe	Xaa	Xaa	Gln
		35					40					45			
Asp	Arg	Ser	Ser	Asn	Ile	Glu	Lys	Ala	Arg	Met	Met	Val	Glu	Val	Glu
	50					55					60				
Ile	Lys	Asp	Gln	Gly	Phe	Arg	Val	Glu	Asp	Ile	Leu	Thr	Glu	Lys	Asn
65					70					75					80
Ile	Gln	Val	Val	Asn	Xaa	Lys	Tyr	Asn	Phe	Xaa	Asn	Glu	Asp	Asp	Leu
				85					90					95	
Phe	Ala	Ala	Val	Gly	Phe	Gly	Gly	Val	Thr	Ser	Leu	Gln	Ile	Val	Asn
			100					105					110		
Lys	Leu	Thr	Glu	Arg	Gln	Arg	Ile	Leu	Asp	Lys	Gln	Arg	Ala	Leu	Asn
		115					120					125			
Glu	Ala	Gln	Glu	Val	Thr	Lys	Ser	Leu	Pro	Ile	Lys	Asp	Asn	Ile	Ile
	130					135					140				
Thr	Asp	Ser	Gly	Val	Tyr	Val	Glu	Gly	Leu	Glu	Asn	Val	Leu	Ile	Lys
145					150					155					160

EP 0 841 394 A2

Leu Xaa Lys Cys Cys Asn Pro Ile Pro Gly Asp Asp Ile Val Gly Tyr  
165 170 175  
5 Ile Thr Lys Gly His Gly Thr Lys Val His Arg Thr Asp Cys Pro Asn  
180 185 190  
Ile Lys Asn Glu Thr Glu Arg Leu Ile Asn Val Glu Trp Val Lys Xaa  
195 200 205  
10 Lys Asp Ala Thr Gln Lys Tyr Gln Val Asp Leu Glu Val Thr Ala Tyr  
210 215 220  
Asp Arg Asn Gly Leu Leu Asn Glu Val Leu Gln Ala Val Ser Ser Thr  
225 230 235 240  
15 Xaa Gly Asn Leu Ile Lys Val Ser Gly Arg Ser Asp Ile Asp Xaa Asn  
245 250 255  
Xaa Ile Ile Asn Ile Ser Val Met Val Lys Asn Val Asn Asp Val Tyr  
260 265 270  
20 Arg Val Val Glu Lys Thr Lys Gln Leu Gly Asp Val Tyr Thr Gly Asn  
275 280 285  
Lys Ser Leu Gly Thr Arg Gly Ser Lys Asn Ile Glu Ser Arg Trp Tyr  
290 295 300  
25 Lys Arg Gly  
305

(2) INFORMATION FOR SEQ ID NO:313:

30 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 92 amino acids  
(B) TYPE: amino acid  
35 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:313:

Val Ser Thr Val His Pro Ile Arg Ile Pro Pro Gly Asp Pro Val Ser  
45 1 5 10 15  
Thr Val Asp Ile Val Ile Gly Arg Val Ala Gln Val His Ile Ala Asp  
20 25 30  
Xaa Val Ile Leu Asp Asn Gly Lys Leu Asp Ile Lys Ser Ile Lys Pro  
50 35 40 45  
Ile Ala Arg Leu Gly Tyr Ser Asp Tyr Thr Val Val Asn Glu Ile Phe  
50 55 60  
Glu Met Lys Ala Pro Xaa Ala Ser Lys Glu Glu Leu Ala Gly Leu Glu  
55 65 70 75 80

EP 0 841 394 A2

Gly Arg Asn Phe Asp Asn Gln Ser Asp Glu Lys Arg  
85 90

(2) INFORMATION FOR SEQ ID NO:314:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 99 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:314:

Val Leu Thr Arg Leu Lys Thr Met Gly Lys Leu Ala Met Leu Met Leu  
1 5 10 15  
Lys Asp Leu Asp Gln Ala Phe Lys Lys Lys Asp Thr Val Leu Ile Arg  
20 25 30  
Glu Ile Ile Glu Arg Asp Glu Asp Ile Asp Asp Leu Tyr Ser His Ile  
35 40 45  
Ile Asn Ala Thr Tyr Leu Ile Asp Asn Asp Pro Phe Val Ala Ala Gln  
50 55 60  
Ala His Leu Ala Ala Arg His Leu Glu Arg Ile Gly Asp His Ile Ile  
65 70 75 80  
Asn Ile Ala Glu Ser Val Tyr Phe Tyr Leu Thr Gly Thr His Tyr Glu  
85 90 95  
Gln Ile Thr

(2) INFORMATION FOR SEQ ID NO:315:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 95 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:315:

Val Thr Asn Lys Ile Leu Ser Gln Phe Leu Asn Ile Lys Pro Pro Ser  
1 5 10 15

EP 0 841 394 A2

Val Ser Glu Met Val Gly Arg Leu Glu Lys Ala Gly Tyr Val Glu Thr  
 20 25 30  
 Lys Pro Tyr Lys Gly Val Arg Leu Thr Glu Asp Gly Leu Thr His Thr  
 5 35 40 45  
 Leu Asp Ile Ile Lys Arg His Arg Leu Leu Glu Leu Phe Leu Ile Glu  
 50 55 60  
 Ile Leu Lys Tyr Asn Trp Glu Glu Val His Gln Glu Ala Glu Ile Leu  
 10 65 70 75 80  
 Glu His Arg Ile Ser Asp Leu Phe Val Glu Arg Leu Asp Ser Cys  
 85 90 95

(2) INFORMATION FOR SEQ ID NO:316:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:316:

Val Ser Gly Asn Xaa Xaa Gln Ala Asp Leu Ile Gly Tyr Ser Phe Lys  
 1 5 10 15  
 Phe Asp Gly Ala Ile Xaa Arg Gln Glu Ala Ser Xaa Asp Val His Ala  
 20 25 30  
 Val Ile Leu Ser Asn Lys Thr Leu  
 35 40

(2) INFORMATION FOR SEQ ID NO:317:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 75 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:317:

Val Val Ser Met Lys Tyr Asn Thr Asn Val Lys His Thr Thr Leu Glu  
 1 5 10 15

EP 0 841 394 A2

Ala Phe Val Thr Thr Val Asn Asp Leu Gly Ile Glu Leu Ile Ile Asn  
 20 25 30  
 Glu Ala Leu Arg Glu Val Arg Lys Arg Gln Leu Ile Glu Leu Ile Asp  
 5 35 40 45  
 Asp Ala Leu Val Asn Lys Asp Glu Ala Ala Phe Asn Gln Tyr Thr Ala  
 50 55 60  
 Glu Tyr Lys Asn Leu Glu Ala Phe Leu Gly Xaa  
 10 65 70 75

(2) INFORMATION FOR SEQ ID NO:318:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 97 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:318:

Val Tyr Lys Val Lys Val Ala Val Ile Met Gly Ser Ser Ser Asp Trp  
 1 5 10 15  
 Lys Ile Met Gln Glu Ser Cys Asn Met Leu Asp Tyr Leu Glu Ile Pro  
 20 25 30  
 Tyr Glu Lys Gln Val Val Ser Ala His Arg Thr Pro Lys Met Met Val  
 35 40 45  
 Gln Phe Ala Ser Glu Ala Arg Glu Arg Gly Ile Asn Ile Ile Ile Ala  
 50 55 60  
 Gly Ala Gly Gly Ala Xaa His Leu Pro Gly Met Val Ala Ser Leu Thr  
 65 70 75 80  
 Thr Leu Pro Val Ile Gly Val Pro Ile Glu Thr Xaa Lys Phe Lys Gly  
 85 90 95  
 Tyr

(2) INFORMATION FOR SEQ ID NO:319:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 42 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

EP 0 841 394 A2

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:319:

Val Lys Lys Ile Asp Cys Ser Xaa Ile Gly Ile Cys Gly Tyr Glu Tyr  
 1 5 10 15  
 Arg Gln Leu Lys Gln Glu Thr Xaa Leu Xaa Leu Phe Xaa Lys Leu His  
 20 25 30  
 Pro Ser Phe Xaa Gly Gln Gly Val Arg Met  
 35 40

(2) INFORMATION FOR SEQ ID NO:320:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 54 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:320:

Val Leu Gly Leu Ile Glu Gly Lys Ser Gln Arg Lys Ala Tyr Ile Asp  
 1 5 10 15  
 Ala Gly Tyr Ser Thr Lys Gly Lys Ser Asp Asn Tyr Ile Asp Ser Arg  
 20 25 30  
 Ala Phe Glu Leu Ser Lys Asn Ser Ala Gly Leu Asp Arg Tyr Glu Glu  
 35 40 45  
 Leu Arg Gln Glu Ala Gly  
 50

(2) INFORMATION FOR SEQ ID NO:321:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 102 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:321:



EP 0 841 394 A2

Val Lys Asp Leu Leu Gln Ala Gln Gln Lys Leu Ile Pro Asp Leu Ile  
 1 5 10 15  
 Asp Lys Met Tyr Lys Arg Phe Ser Ile Leu Thr Thr Ile Ser Lys Asn  
 5 20 25 30  
 Gln Pro Val Gly Arg Arg Ser Leu Ser Glu His Met Asp Met Thr Glu  
 35 40 45  
 Arg Val Leu Arg Ser Glu Thr Asp Met Leu Lys Lys Gln Asp Val Ile  
 10 50 55 60  
 Lys Val Lys Pro Thr Gly Met Glu Ile Thr Ala Glu Gly Glu Gln Leu  
 65 70 75 80  
 Ile Ser Gln Leu Lys Gly Tyr Xaa Asp Ile Tyr Gly Asp Asp Asn Arg  
 15 85 90 95  
 Leu Val Lys Lys Gly Ile  
 100

(2) INFORMATION FOR SEQ ID NO:322:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:322:

Val Met Ile Met Arg Lys Glu Ile Glu Ala Leu Ile Phe Ser Asp Val  
 1 5 10 15  
 Ser Ser Tyr Asp Ile Tyr Val Asn Thr Gly Val Asn Gln Gly Leu Ile  
 20 25 30  
 Gly Asp Ile Lys Asp Gly Tyr Leu Thr Ile Asp Ser Met Pro Tyr Ile  
 35 40 45  
 Asp Ala Glu Arg Leu Tyr His Phe Ala Tyr Gly Thr  
 45 50 55 60

(2) INFORMATION FOR SEQ ID NO:323:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 281 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

EP 0 841 394 A2

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:323:

```

Val Thr Cys Ile Val Leu Val Lys Val Val Glu Asn Ala Glu Ile Lys
  1           5           10           15
10 Gly Ile Glu Ile Phe Asn Ser Val Asp Asn Asn Asn Lys Lys Ile Phe
    20           25           30
    Leu Lys Asp Gly Asn Val Val Gly Ala Val Ser Cys Gln Tyr Gly Asp
    35           40           45
15 Ile Asp Asp Gly Ser Arg Phe Tyr Asn Met Met Lys Lys Gly Glu Ser
    50           55           60
    Thr Glu Asp Tyr Thr Leu Val Ser Leu Leu Thr Lys Gly Gly Glu Glu
    65           70           75           80
20 Ala Ser Leu Ser Ile Ala Asp Met Ala Asp Asp Glu Thr Ile Cys Gly
    85           90           95
    Cys Asn Gly Val Asp Lys Gly Thr Ile Val Asn Ala Ile Thr Glu Asn
    100          105          110
25 Gly Phe Thr Thr Val Glu Glu Val Thr Ala Lys Thr Lys Ala Gly Asn
    115          120          125
    Ser Cys Gly Lys Cys Lys Pro Gln Ile Ala Gln Ile Leu Gln His Thr
    130          135          140
30 Leu Gly Asp Asp Phe Val Ala Ala Lys Pro Ala Gly Ile Cys Gly Cys
    145          150          155          160
    Thr Asp Leu Thr Arg Asp Gln Ile Val Thr Gln Ile Arg Ala Lys Gly
    165          170          175
35 Leu Lys Thr Ser Lys Glu Val Arg His Val Leu Asn Phe Lys Asn Lys
    180          185          190
    Gly Gly Cys Pro Lys Cys Arg Pro Ala Ile Asn Tyr Tyr Leu Asn Met
    195          200          205
40 Val Tyr Pro His Asp His Glu Asp Glu Arg Glu Ser Arg Phe Ala Asn
    210          215          220
    Glu Arg Tyr His Ala Asn Ile Gln Asn Asp Gly Thr Phe Ser Val Ile
    225          230          235          240
45 Pro Gln Met Arg Gly Gly Val Thr Asp Ala Asp Gln Leu Ile Arg Leu
    245          250          255
    Gly Glu Val Ala Lys Lys Tyr His Val Pro Leu Val Lys Val Thr Gly
    260          265          270
50 Ser Gln Arg Val Gly Leu Tyr Gly Gly
    275          280

```

(2) INFORMATION FOR SEQ ID NO:324:

(i) SEQUENCE CHARACTERISTICS:

EP 0 841 394 A2

(A) LENGTH: 70 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:324:

Val	Pro	Met	Thr	Pro	Leu	Val	Glu	Pro	Xaa	Pro	Xaa	Thr	Val	Ser	Ser
1				5				10					15		
Asn	Asn	Lys	Ser	Asn	Ser	Asn	Ser	Ser	Thr	Leu	Xaa	Tyr	Leu	Arg	Thr
			20					25					30		
Leu	Glu	Asn	Arg	Gly	Trp	Asp	Phe	Asp	Gly	Ser	Tyr	Gly	Trp	Gln	Cys
			35				40					45			
Phe	Asp	Leu	Xaa	Asn	Val	Phe	Trp	Xaa	Ser	Ser	Ser	Met	Gly	Met	Asp
			50				55					60			
Leu	Lys	Gly	Phe	Trp	Arg										
65							70								

(2) INFORMATION FOR SEQ ID NO:325:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 122 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:325:

Val	Leu	Gln	Leu	Ile	Thr	Gln	Ile	Tyr	Gly	Ala	Glu	Lys	Arg	Val	Xaa
1				5				10					15		
Trp	Val	Asn	Gln	Leu	Val	Ser	Xaa	Ser	Phe	Asn	Lys	Lys	Ile	Arg	Pro
			20					25					30		
Leu	Arg	Thr	Xaa	Ser	Asp	Trp	Ile	Ser	Ser	Asn	Pro	Ile	Glu	Val	Asp
			35				40					45			
Asn	Phe	Ile	Lys	Asp	Pro	Tyr	Ser	Gly	Phe	Asn	Val	Ser	Xaa	Gln	Leu
			50				55					60			
Leu	Tyr	His	Thr	Ala	Tyr	Tyr	Met	Leu	His	Thr	Ser	Gln	Leu	Lys	Asn
65				70				75				80			
Met	Lys	Met	Leu	Asn	His	Ala	Met	Pro	Ile	Leu	Leu	Val	Ser	Gly	Tyr
				85				90						95	

EP 0 841 394 A2

Ala Asp Ser Leu Gly Asp Tyr Gly Lys Gly Ile Leu Lys Leu Ala Asn  
 100 105 110  
 5 Ile Tyr Arg Glu Ala Gly Ile Lys Asn Cys  
 115 120

(2) INFORMATION FOR SEQ ID NO:326:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 56 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:326:

Val Thr Leu Glu Val Ala Glu Met Xaa Asn Ser Xaa Lys Asp Xaa Lys  
 1 5 10 15  
 Lys Phe Glu Ile Ile Asp Arg Xaa Lys Ser Phe Tyr Asp Glu Glu Gln  
 20 25 30  
 Xaa Ile Asp Leu Val Phe Val Val Asn Gln Ile Asn Gly Trp Asn Arg  
 35 40 45  
 30 Leu Asn Ile Ile Ser Asp Arg Leu  
 50 55

(2) INFORMATION FOR SEQ ID NO:327:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 119 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:327:

Val Asp Ser Val Thr Ile Lys Val Leu Asn Glu Pro Ser Pro Lys Leu  
 1 5 10 15  
 Leu Thr Thr Trp Tyr Ala Glu Gln Val Thr Gln Gly Lys Ile Lys Thr  
 20 25 30  
 55 Ser Lys Tyr Val Lys Lys Glu Cys Glu Arg His Leu Arg Tyr Leu Glu  
 35 40 45

EP 0 841 394 A2

Asn Gly Gly Lys Trp Val Phe Asp Glu Glu Leu Ala His Arg Pro Ile  
 50 55 60  
 5 Arg Phe Ile Glu Lys Phe Cys Lys Xaa Ser Lys Gly Ser Lys Arg Gln  
 65 70 75 80  
 Leu Ala Leu Gln Pro Trp Gln His Phe Ile Ile Gly Ser Leu Phe Gly  
 85 90 95  
 10 Trp Val His Lys Glu Thr Lys Leu Arg Arg Phe Lys Glu Ala Leu Ile  
 100 105 110  
 Phe Met Gly Ala Lys Lys Trp  
 115

(2) INFORMATION FOR SEQ ID NO:328:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:328:

Val Ile Xaa Xaa Val Ile Lys Glu Lys Pro Xaa Ala Arg Trp Leu Phe  
 1 5 10 15  
 Ser Pro Xaa Ser Pro Xaa Xaa Ala Ile Asp Gly Asp Thr Leu Val Thr  
 20 25 30  
 Lys Xaa Xaa Ala Ser Asn  
 35

(2) INFORMATION FOR SEQ ID NO:329:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 211 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:329:

Val Leu Asn Leu Ser Ile Ser Leu Thr Ser Leu Ala Thr Gly Ile Phe  
 1 5 10 15

EP 0 841 394 A2

```

... Met Val Ala Ala Gly Asp Ile Ala Asp Lys Ile Gly Gln Xaa Arg Met...
      20          25          30
5   Thr Tyr Met Gly Leu Ile Ile Xaa Met Phe Xaa Ser Leu Leu Leu Ile
      35          40          45
      Ile Ser Asp Ile Thr Ala Leu Leu Ile Ile Xaa Arg Asn Phe Thr Arg
      50          55          60
10  Ser Ile Ser Xaa Tyr Leu Gly Tyr Pro Pro Thr Val Gly Val Leu Asn
      65          70          75          80
      Asn Gln Phe Lys Gly Glu His Leu Arg Arg Ala Ile Ser Tyr Leu Met
      85          90          95
15  Ile Ser Thr Val Gly Gly Ile Gly Leu Ala Gly Val Ile Gly Gly Leu
      100         105         110
      Ile Ala Thr Asn Phe Gly Trp Gln Met Asn Phe Ile Ile Ser Ile Val
      115         120         125
20  Ile Xaa Phe Ile Ala Ile Leu Leu Leu Lys Gly Thr Pro Glu Lys Val
      130         135         140
      Ser Gln His Xaa His Arg His Pro Phe Asp Tyr Lys Gly Met Ser Ile
      145         150         155         160
25  Phe Ala Val Met Ile Gly Ser Phe Thr Leu Leu Leu Thr Gln Gly Phe
      165         170         175
      Glu Gln Gly Trp Phe Ser Thr Phe Ser Phe Ile Cys Leu Ser Ile Phe
      180         185         190
30  Ile Ile Thr Thr Leu Ile Phe Ile Ile Ile Glu Arg Arg His Glu Val
      195         200         205
      Pro Phe Tyr
      210

```

35

(2) INFORMATION FOR SEQ ID NO:330:

(i) SEQUENCE CHARACTERISTICS:

40

- (A) LENGTH: 44 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

45

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:330:

50

```

Val Ala Pro Phe Leu Tyr Leu Xaa Ile Asn Asp Glu Xaa Xaa Asp Met
  1          5          10          15
Lys Ile Ala Gly Asp Leu Asn Ala Tyr Asn Ala Leu Xaa Ala Tyr Thr
55          20          25          30

```

EP 0 841 394 A2

Val Xaa Arg Glu Leu Gly Leu Asn Glu Gln Xaa Asn  
35 40

(2) INFORMATION FOR SEQ ID NO:331:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:331:

Val Tyr Val Ile Ser Leu Asn Asp Asn Ala Ala Asp Gly Arg Asp Thr  
1 5 10 15  
Ser Trp Ile Tyr Asp Ala Asp Phe Gly Lys Ile Ile  
20 25

(2) INFORMATION FOR SEQ ID NO:332:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 56 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:332:

Val Met Ser Lys Leu Arg Pro Leu Leu His Tyr Ala Arg Ala Lys Asp  
1 5 10 15  
Ile Asp Asn Tyr Arg Thr Val Glu Glu Ser Tyr Arg Gln Gly Gln Tyr  
20 25 30  
Tyr Arg Xaa Xaa Ile Val Asp Gly Lys Leu Asn Ile Gln Phe Asn Glu  
35 40 45  
Gly Glu Pro Tyr Phe Gly Arg His  
50 55

(2) INFORMATION FOR SEQ ID NO:333:

(i) SEQUENCE CHARACTERISTICS:

EP 0 841 394 A2

(A) LENGTH: 129 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:333:

```

Val His Pro Asp Val Leu Leu Thr Arg Tyr Val Glu Gly Lys Glu Asp
 1           5           10           15
Gln Val Glu Lys Val Leu Tyr Gln Leu Ala Asp Ile Asp Ile Ser Glu
 20           25           30
Ile Pro Lys Asp Phe Ile Leu Leu Pro Thr His Pro Tyr Gln Xaa Asn
 35           40           45
Val Leu Xaa Gln Tyr Pro Gln Tyr Met Gln Tyr Ser Glu Gln Gly Leu
 50           55           60
Ile Lys Asp Leu Gly Val Ser Gly Asp Leu Val Ser Pro Thr Ser Xaa
 65           70           75           80
Xaa Arg Thr Val Phe Ser Lys Ala Leu Asn Ile Tyr Leu Xaa Xaa Pro
 85           90           95
Ile His Val Xaa Ile Thr Asn Phe Xaa Arg Thr Asn Asp Leu Glu Gln
100           105           110
Ile Xaa Arg Thr Ser Cys Met Pro Arg Met Ser Ser Ser Ser Asp Gln
115           120           125
Arg

```

(2) INFORMATION FOR SEQ ID NO:334:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 113 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:334:

```

Val Gln Asn Lys Glu Leu Ile Gln His Ala Ala Tyr Ala Ala Ile Glu
 1           5           10           15
Arg Ile Leu Asn Glu Tyr Phe Arg Glu Glu Asn Leu Tyr Gln Val Pro
 20           25           30

```



EP 0 841 394 A2

.. Pro Gln Asn His Gln Trp Ser Ile Gln Leu Ser Glu Leu Glu Thr Leu ..  
 35 40 45  
 5 Thr Gly Glu Phe Arg Tyr Trp Ser Ala Met Gly His His Met Tyr His  
 50 55 60  
 Pro Glu Val Trp Leu Ile Asp Gly Lys Ser Lys Lys Ile Thr Thr Tyr  
 65 70 75 80  
 10 Lys Glu Ala Ile Ala Arg Ile Leu Pro His Met Ala Gln Ser Ala Asp  
 85 90 95  
 Asn Gln Thr Ala Val Gln Pro Thr Tyr Gly Xaa Asn Tyr Val Xaa His  
 100 105 110  
 15 Arg

(2) INFORMATION FOR SEQ ID NO:335:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:335:

Val Ala Val Ser Phe Leu Thr Ser Arg Asp Ile Pro Xaa Asp Lys Val  
 1 5 10 15  
 35 Val Lys Met Asp Val Asp Pro Pro Xaa Tyr Leu  
 20 25

(2) INFORMATION FOR SEQ ID NO:336:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 69 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:336:

Val Lys Asn Val Lys Thr Gln Ser Ile Gln Asn Ala Ser Ile Val Thr  
 1 5 10 15

# EP 0 841 394 A2

Val Glu Tyr Glu Asn Asn Thr Asp Met Asp Lys Ala Glu Glu Gln Leu  
 20 25 30  
 Lys Lys Glu Ile Asp Lys Ile Lys Phe Lys Asp Glu Val Gly Gln Pro  
 35 40 45  
 Glu Leu Arg Arg Asn Ser Met Asp Ala Phe Pro Val Leu Ala Tyr Ser  
 50 55 60  
 Phe Gln Ile Lys Arg  
 65

## (2) INFORMATION FOR SEQ ID NO:337:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 180 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Protein

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:337:

Val Met His Xaa Ile Phe Pro Val Val Ser Xaa Thr Phe Ala Ser Phe  
 1 5 10 15  
 Lys Ser Met Tyr Gly Gly Ile Pro Gln Asp Phe Ile Asp Tyr Leu Phe  
 20 25 30  
 Ile Asp Glu Ala Gly Gln Ala Ile Pro Gln Ala Ala Val Gly Ala Leu  
 35 40 45  
 Tyr Arg Ser Lys Lys Val Val Ala Val Gly Asp Pro Ile Gln Ile Glu  
 50 55 60  
 Pro Val Val Thr Leu Glu Ser His Leu Ile Asp Asn Ile Arg Lys Asn  
 65 70 75 80  
 Tyr His Val Pro Glu Tyr Leu Val Ser Lys Glu Ala Ser Val Gln Ser  
 85 90 95  
 Val Ala Asp Asn Ala Asn Gln Tyr Gly Phe Trp Lys Ser Asp Ala Thr  
 100 105 110  
 Asp Ser Asn Gln Lys Thr Trp Ile Gly Ile Pro Leu Trp Val His Arg  
 115 120 125  
 Arg Cys Leu Lys Pro Met Phe Thr Ile Ala Asn Gln Ile Ala Tyr Asn  
 130 135 140  
 Asn Lys Met Val Leu Ala Ser Asn Ile Thr Lys Val Gly Lys Thr Gly  
 145 150 155 160  
 Trp Tyr Asp Val Lys Gly Asn Ala Val Gln Asn Ile Xaa Glu Arg Ala  
 165 170 175

EP 0 841 394 A2

... Trp Val Lys Lys  
180

5

(2) INFORMATION FOR SEQ ID NO:338:

(i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 112 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:338:

20

Val	Lys	Phe	Phe	Asp	Tyr	Thr	Phe	Ala	Ser	Val	Gly	Val	Lys	Pro	Asn
1				5					10					15	
Glu	Leu	Lys	Gln	Phe	Asp	Tyr	Lys	Met	Val	Glu	Val	Thr	Gln	Gly	Ala
			20					25					30		
His	Ala	Asn	Tyr	Tyr	Pro	Gly	Asn	Ser	Pro	Leu	His	Leu	Arg	Val	Tyr
		35					40					45			
Tyr	Asp	Thr	Ser	Asn	Arg	Gln	Ile	Leu	Arg	Ala	Ala	Ala	Val	Gly	Lys
	50					55				60					
Glu	Gly	Ala	Asp	Lys	Arg	Ile	Asp	Val	Leu	Ser	Met	Ala	Met	Met	Asn
65					70					75				80	
Gln	Leu	Thr	Val	Asp	Glu	Leu	Thr	Glu	Phe	Xaa	Val	Xaa	Tyr	Ala	Pro
			85					90					95		
Pro	Tyr	Trp	His	Pro	Lys	Asp	Leu	Xaa	Asn	Met	Ile	Gly	Leu	Gln	Ser
			100					105					110		

40

(2) INFORMATION FOR SEQ ID NO:339:

(i) SEQUENCE CHARACTERISTICS:

45

- (A) LENGTH: 97 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

50

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:339:

55

Val	Met	Gln	Ser	Ser	Lys	Trp	Asn	Ala	Met	Ser	Leu	Leu	Met	Asp	Glu
1					5				10					15	

# EP 0 841 394 A2

Lys Thr Lys Gln Ala Glu Val Leu Arg Thr Ala Ile Asp Glu Ala Asp  
 20 25 30  
 5 Ala Ile Val Ile Gly Ile Gly Ala Gly Met Ser Ala Ser Asp Gly Phe  
 35 40 45  
 Thr Tyr Val Gly Glu Arg Phe Thr Glu Asn Phe Pro Asp Phe Ile Glu  
 50 55 60  
 10 Lys Tyr Arg Phe Phe Asp Met Leu Gln Ala Ser Leu His Pro Tyr Gly  
 65 70 75 80  
 Ser Trp Gln Glu Tyr Trp Ala Phe Glu Ser Arg Phe Ile Thr Xaa Asn  
 85 90 95  
 15 Tyr

## (2) INFORMATION FOR SEQ ID NO:340:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 93 amino acids  
 (B) TYPE: amino acid  
 25 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:340:

Val Tyr Pro Ile Lys Ala Ile Ile Asp Ser Leu Thr Thr Phe Ser Gly  
 1 5 10 15  
 35 Ile Glu His Arg Leu Gln Tyr Val Gly Thr Asn Arg Thr Asn Lys Tyr  
 20 25 30  
 Tyr Asn Asp Ser Xaa Ala Thr Asn Thr Leu Ala Thr Gln Phe Ala Leu  
 35 40 45  
 40 Asn Ser Phe Asn Gln Pro Ile Ile Trp Leu Leu Trp Trp Tyr Trp Ile  
 50 55 60  
 Glu Gly Asp Glu Phe Asp Glu Leu Ile Pro Tyr Met Glu Asn Val Xaa  
 45 65 70 75 80  
 His Asp Gly Cys Ile Arg Thr Asn Glu Ser Xaa Val Cys  
 85 90

## 50 (2) INFORMATION FOR SEQ ID NO:341:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 81 amino acids  
 55 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single

EP 0 841 394 A2

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:341:

Val Met Gly Thr Asp Lys Arg Val Ser Pro Asp Leu Phe Ile Lys Arg  
 1 5 10 15  
 Leu Asn Phe Tyr Asp Tyr Leu Leu Leu Asn Ser Asp Gly Leu Thr Asp  
 20 25 30  
 Tyr Val Lys Asp Asn Glu Ile Lys Arg Leu Leu Val Lys Glu Gly Thr  
 35 40 45  
 Ile Glu Asp His Gly Asp Gln Leu Met Gln Leu Ala Leu Asp Asn His  
 50 55 60  
 Ser Lys Asp Asn Val Thr Phe Ile Leu Ala Ala Ile Xaa Gly Asp Lys  
 65 70 75 80  
 Val

(2) INFORMATION FOR SEQ ID NO:342:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:342:

Val Lys Leu Lys Gly Met Arg Lys Gln Tyr Glu Val Gly Glu Gln Ala  
 1 5 10 15  
 Ser Pro Ile Glu Phe Val Thr Gly Arg Val Asn Pro His Arg  
 20 25 30

(2) INFORMATION FOR SEQ ID NO:343:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 82 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

EP 0 841 394 A2

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:343:

```

Val Glu Val Gln Val Met Lys Phe Gly Lys Thr Ile Ala Val Val Leu
  1             5             10             15
Ala Ser Ser Val Leu Leu Xaa Gly Cys Thr Thr Asp Lys Lys Glu Ile
  20             25             30
Lys Ala Tyr Leu Lys Gln Val Asp Lys Ile Lys Asp Asp Glu Glu Pro
  35             40             45
Ile Xaa Thr Val Xaa Lys Lys Ile Ala Glu Leu Asp Glu Lys Lys Lys
  50             55             60
Lys Leu Thr Glu Asp Val Asn Ser Lys Asp Thr Ala Gly Ser Arg Xaa
  65             70             75             80
Lys Gln

```

(2) INFORMATION FOR SEQ ID NO:344:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 162 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:344:

```

Val Ile Gln Ser Ser Ser Leu Leu Asn Xaa Phe Lys Val Phe Ala Ala
  1             5             10             15
Ile Ser Asp Thr Val Ile His Gly Ser Asp Val Asp Leu Val Asn Trp
  20             25             30
Val Glu Asn Ala Leu Gln Ser Gly Leu Ala Ala Lys Glu Arg Lys Asp
  35             40             45
Ala Leu Ala Val Ser Pro Lys Ser Gly Leu Glu His Met Ser Phe Tyr
  50             55             60
Asn Gly Ser Val Lys Glu Val Xaa Gly Ala Asp Met Tyr Asp Ala Ile
  65             70             75             80
Ile Asn Ala Leu Ala Asp Leu His Glu Asp Tyr Arg Asp Asn Ala Thr
  85             90             95
Ile Tyr Met Arg Tyr Ala Asp Tyr Val Lys Ile Ile Ser Val Leu Ser
 100             105             110

```

# EP 0 841 394 A2

Asn Gly Thr Thr Asn Phe Phe Asp Thr Pro Ala Glu Lys Val Xaa Gly  
 115 120 125  
 5 Xaa Pro Val Gly Ile Tyr Arg Cys Arg Xaa Leu Pro Leu Leu Trp Glu  
 130 135 140  
 Asp Phe Asn Tyr Cys Gly Xaa Tyr Leu Leu Met Gly Gln Pro Tyr Asp  
 145 150 155 160  
 10 Thr Asp

## (2) INFORMATION FOR SEQ ID NO:345:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Protein

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:345:

Val Tyr Lys Thr Asp Leu Tyr Xaa Gln Xaa Xaa Lys Ile Asp Thr Val  
 1 5 10 15  
 30 Lys Val Ile Lys Glu Val Glu Leu Lys Lys Pro Ile Thr Tyr Glu Ala  
 20 25 30  
 Gly Ala Thr Ser Asp Ser Lys Leu Xaa Lys Glu Trp Met Asp Ser  
 35 35 40 45

## (2) INFORMATION FOR SEQ ID NO:346:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 62 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Protein

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:346:

Val Asn Pro Arg Tyr Lys Xaa Thr Leu Leu His Arg Arg Leu His Thr  
 1 5 10 15  
 55 Ile Leu Xaa Lys Met Leu Glu Asp Leu Ser Phe Glu Ala Pro Ser Met  
 20 25 30

EP 0 841 394 A2

Pro Asn Ala Val Val Asp Ile Thr Pro Gln Tyr Val Asp Asp Lys Leu  
 35 40 45  
 Lys Ser Ile Ser Thr Xaa Lys Asp Leu Ser Ala Phe Ile Leu  
 5 50 55 60

(2) INFORMATION FOR SEQ ID NO:347:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 117 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:347:

Val Lys Ile Leu Lys Asp Thr Asn Leu Met Tyr Leu Glu Trp Asn Tyr  
 1 5 10 15  
 Leu Val Xaa Arg Val Val Ala Val Val Glu Xaa Gln Gly Thr Leu Ser  
 20 25 30  
 Asp Ala Val Asn Lys Ala Leu Gln Tyr Trp Val Ser His Val Asp Asp  
 35 40 45  
 Thr His Tyr Leu Leu Gly Ser Ala Leu Gly Pro Asp Pro Phe Pro Thr  
 50 55 60  
 Ile Val Arg Asp Phe Gln Ser Val Ile Gly Lys Glu Xaa Lys Ser Gln  
 65 70 75 80  
 Ile Leu Lys Lys Glu Gly Arg Leu Pro Asp Ala Ile Val Ala Cys Ile  
 85 90 95  
 Gly Gly Gly Ser Asn Ala Ile Gly Leu Val Pro Phe Ile His Leu Leu  
 100 105 110  
 Lys Met Met Leu His  
 115

(2) INFORMATION FOR SEQ ID NO:348:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 135 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein



EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:348:

```

5   Val Gly Arg Leu Pro Lys Lys Ile Gln Asn Ile Glu Gln Arg Pro Xaa
    1           5           10           15
    Gly Val Ser Xaa Thr Thr Phe Thr Ile Ala Val Asn Arg Thr Phe Thr
      20           25           30
10  Asn Ala Gln Gly Glu Arg Glu Ala Asp Phe Ile Asn Cys Val Thr Phe
      35           40           45
    Arg Lys Gln Ala Glu Asn Val Asn Asn Tyr Leu Ser Lys Gly Ser Leu
      50           55           60
15  Ala Gly Val Asp Gly Arg Leu Gln Ser Arg Ser Tyr Asp Asn Lys Glu
      65           70           75           80
    Gly Arg Arg Val Phe Val Thr Glu Val Val Ala Asp Ser Val Gln Phe
      85           90           95
20  Leu Glu Pro Lys Asn Asn Asn Gln Xaa Gln Asn Asn Asn Tyr Gln Gln
      100          105          110
    Gln Gly Gln Ala Gln Thr Gly Asn Asn Pro Phe Asp Asn Ser Glu Glu
      115          120          125
25  Asp Phe Ser Asp Leu Pro Phe
      130          135

```

(2) INFORMATION FOR SEQ ID NO:349:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 146 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:349:

```

45  Val Val Ser Gln Ser Ile Leu Leu Thr Leu Ser Thr Xaa Leu Thr Asn
    1           5           10           15
    Val Thr Xaa Leu Thr Ser Xaa Ala Gln Xaa Pro Ala Asp Tyr Val Lys
      20           25           30
    Asn Met Ile Thr Gly Ala Ala Gln Met Asp Gly Gly Ile Leu Val Val
      35           40           45
50  Ser Ala Ala Asp Gly Pro Met Pro Gln Thr Arg Glu His Ile Leu Leu
      50           55           60
    Ser Arg Asn Val Gly Val Pro Ala Leu Val Val Phe Leu Asn Lys Val
      65           70           75           80

```

# EP 0 841 394 A2

Asp Met Val Asp Asp Glu Glu Leu Leu Glu Leu Val Glu Met Glu Val  
 85 90 95  
 Arg Asp Leu Leu Ser Glu Tyr Asp Phe Pro Gly Asp Asp Val Pro Val  
 100 105 110  
 Ile Ala Gly Ser Ala Leu Lys Ala Leu Glu Gly Asp Ala Gln Tyr Glu  
 115 120 125  
 Glu Lys Ile Leu Glu Leu Met Glu Ala Val Asp Thr Tyr Ile Pro Leu  
 130 135 140  
 Gln Thr  
 145

(2) INFORMATION FOR SEQ ID NO:350:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 65 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:350:

Val Lys Val Ser Gly Cys Ser Asp Leu Val Gln Cys Ala Ile His Asp  
 1 5 10 15  
 Cys Gln Tyr Tyr Glu Val Gln Met Pro Gln Ala His Arg Val Val Pro  
 20 25 30  
 Thr Thr Ile Asn Met Gly Asn His Ser Trp Lys Ala Leu Glu Ile Ile  
 35 40 45  
 Tyr Glu Thr Tyr Cys Thr Leu Cys Asp Ser Xaa Leu Asn Lys Xaa Thr  
 50 55 60  
 Thr  
 65

(2) INFORMATION FOR SEQ ID NO:351:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 67 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:351:

```

5   Val Xaa Val Asn Ile Asn Cys Ala Asn Pro Glu Cys Asn Lys Gln Ile
    1           5           10           15
    Leu Val Ser Glu Glu Xaa Glu Thr Lys Tyr Leu Gly Ala Cys Ser Tyr
      20           25           30
10  Glu Cys Ala Lys His Glu Arg Asn Arg Tyr Val Gln Ala Asn Asn Ile
      35           40           45
    Ser Asp Asn Glu Trp Gln Gln Arg Leu Thr Asn Phe Asp Asp Leu His
      50           55           60
15  Gln His Ala
    65

```

(2) INFORMATION FOR SEQ ID NO:352:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 63 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:352:

```

35  Val Thr Arg Ser Lys Met Ile Val Val Glu Ser Tyr Asp Val Asn Gly
    1           5           10           15
    Asn Lys Val Glu Leu Thr Ala His Glu Asp Val Ala Arg Met Ile Leu
      20           25           30
    His Ile Ile Asp Gln Met Asn Gly Leu Pro Phe Xaa Glu Arg Ala Asp
      35           40           45
40  Arg Ile Leu Thr Asp Lys Glu Val Glu Ala Tyr Phe Ile Asn Asp
      50           55           60

```

(2) INFORMATION FOR SEQ ID NO:353:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 40 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:353:

5 Val Thr Lys Asn Ala Gly Asn His Glu Ile Glu Ile Asp Arg Pro Ile  
1 5 10 15  
Val Asn Pro Ile Lys Lys Leu Pro Thr Arg Phe Thr Ile Asn Val Ala  
20 25 30  
10 Ile Gly Asn Phe Glu Val Asn Leu  
35 40

(2) INFORMATION FOR SEQ ID NO:354:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 111 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:354:

30 Val Lys Gly Gln Asn Leu Thr His Gly Thr Asn Ala Ile Asp Tyr Ile  
1 5 10 15  
Thr Phe Asp Pro Asn Thr Asn Thr Asn Gly Ile Thr Ala Ala Trp Ala  
20 25 30  
Asn Arg Gln Gln Pro Asn Asn Gln Gln Ala Gly Val Gln His Leu Asn  
35 35 40 45  
Val Asp Val Thr Tyr Pro Gly Ile Ser Ala Ala Lys Arg Val Pro Val  
50 55 60  
Thr Val Asn Val Tyr Gln Phe Gly Ile Pro Ser Asn Tyr Leu Tyr Asn  
40 65 70 75 80  
Asn Ser Trp Trp His Phe Ser Lys Trp Tyr Ala Ser Ile Arg Ile Cys  
85 90 95  
Thr Tyr Ala Lys Arg Phe Arg Phe Thr Asn Arg Trp Ile Tyr Val  
45 100 105 110

(2) INFORMATION FOR SEQ ID NO:355:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 152 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

EP 0 841 394 A2

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:355:

```

Val Thr Thr Gln Asn Gln Ala Ile Asp Asn Thr Thr Gly Ala Thr Thr
 1           5           10           15
10 Glu Glu Lys Asn Ala Ala Lys Asp Leu Val Leu Xaa Ala Lys Glu Lys
    20           25           30
Ala Xaa Gln Asp Ile Leu Xaa Ala Gln Xaa Thr Asn Asp Val Thr Gln
    35           40           45
15 Ile Asn Asp Gln Ala Val Ala Asp Ile Gln Gly Ile Thr Gly Asp Thr
    50           55           60
Thr Ile Xaa Asp Val Ala Ile Asp Xaa Xaa Ala Thr Pro Ala Thr Xaa
65           70           75           80
20 Xaa Xaa Ala Leu Xaa Gly His Xaa Ala Xaa Ala Thr Xaa Glu Xaa Xaa
    85           90           95
Xaa Gln Xaa Xaa Xaa Thr Lys Xaa Xaa Ala Gln Xaa Xaa Thr Xaa Val
    100          105          110
25 Phe Thr Xaa Leu Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Val Xaa
    115          120          125
Xaa Leu Ala Xaa Xaa Xaa Arg Thr Leu Xaa Ser Arg Xaa Xaa Xaa His
    130          135          140
30 Xaa Xaa Xaa Xaa Xaa Asn Arg Phe
    145          150

```

(2) INFORMATION FOR SEQ ID NO:356:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 65 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:356:

```

Val Val Ser Leu Ile Asp Asp Glu Asp Asp Asn Xaa Asn Asp Arg Gln
50  1           5           10           15
Leu Val Val Ser Ala Pro Ser Lys Lys Pro Thr Thr Pro Thr Thr Tyr
    20           25           30
Thr Glu Thr Thr Thr Gln Val Xaa Met Pro Thr Val Glu Arg Gln Thr
55  35           40           45

```

# EP 0 841 394 A2

Gln Gln Gln Ile Val Tyr Lys Thr Pro Lys Thr Ile Ser Trp Ile Lys...  
 50 55 60

5 Trp  
 65

(2) INFORMATION FOR SEQ ID NO:357:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 120 amino acids

(B) TYPE: amino acid

15

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:357:

Val Lys Ile Met Glu Met Ser Val Thr Glu Val Ile Phe Ser Phe Leu  
 1 5 10 15  
 25 Gly Gly Leu Gly Ile Phe Leu Tyr Gly Leu Lys Ile Met Gly Asp Gly  
 20 25 30  
 Leu Gln Ala Ser Ala Gly Asp Arg Leu Arg Asp Ile Leu Asn Lys Phe  
 35 40 45  
 30 Thr Ser Asn Pro Val Leu Gly Val Ile Ala Gly Ile Val Val Thr Ile  
 50 55 60  
 Leu Ile Gln Ser Ser Ser Gly Thr Thr Val Ile Thr Ile Gly Leu Val  
 35 65 70 75 80  
 Thr Ala Gly Phe Met Thr Leu Lys Gln Ala Ile Gly Val Ile Met Gly  
 85 90 95  
 Ala Asn Ile Gly Thr Thr Val Thr Ala Phe Ile Ile Xaa Ile Arg Phe  
 40 100 105 110  
 Arg Arg Ile Cys Lys Cys Gln Phe  
 115 120

45

(2) INFORMATION FOR SEQ ID NO:358:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 40 amino acids

50

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

55

(ii) MOLECULE TYPE: Protein

EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:358:

5 Val Leu Cys Asp Ala Asp Val Leu Ala Leu Ile Asp Ile Asp Val Asp  
 1 5 10 15  
 Ile Asp Val Leu Lys Glu Pro Asp Ala Leu Val Leu Val Asp Cys Leu  
 20 25 30  
 10 Val Thr Leu Asn His Xaa Met Met  
 35 40

(2) INFORMATION FOR SEQ ID NO:359:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 68 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:359:

30 Val Leu Val Glu Leu Glu Ser Leu Thr Leu Val Glu Pro Asp Ile Glu  
 1 5 10 15  
 Ser Leu Lys Leu Val Asp Val Glu Thr Asp Thr Leu Pro Leu Ile Glu  
 20 25 30  
 Ser Asp Val Glu Ser Asp Val Leu Val Glu Phe Asp Pro Leu Met Leu  
 35 35 40 45  
 Asp Glu Ser Leu Val Asp Ile Glu Ser Leu Ser Asp Ala Leu Met Leu  
 50 55 60  
 40 Ile Glu Ser Asn  
 65

(2) INFORMATION FOR SEQ ID NO:360:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 115 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:360:

EP 0 841 394 A2

Val Ile Ser Lys Xaa Gln Trp Val Val Val His Phe Tyr Pro Ser Trp  
 1 5 10 15  
 Lys Asp Ala Glu Gly His Phe Glu Asn Asn Gly Arg Cys Asn Leu Gly  
 5 20 25 30  
 Val Xaa Thr Leu Asn Leu Pro Arg Met Ala Leu Glu Ser Ala Gly Asn  
 35 40 45  
 Met Thr Lys Phe Trp Glu Ile Phe Tyr Glu Arg Ile Asp Val Leu His  
 10 50 55 60  
 Asp Ala Leu Leu Tyr Arg Ile Asn Arg Leu Lys Asp Xaa Val Pro Asn  
 65 70 75 80  
 Asn Ala Pro Ile Leu Tyr Lys Ser Gly Ala Phe Asn Tyr Lys Leu Lys  
 15 85 90 95  
 Glu Thr Asp Asp Val Ala Glu Leu Phe Lys Asn Lys Arg Ala Thr Ile  
 100 105 110  
 Ser Met Gly  
 20 115

(2) INFORMATION FOR SEQ ID NO:361:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 61 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:361:

Val Xaa Ser Thr Lys Thr Ala Leu Xaa Gly Asp Ala Lys Leu Asn Glu  
 1 5 10 15  
 Ala Lys Ala Ala Ala Lys Gln Thr Leu Gly Thr Leu Thr His Ile Asn  
 20 25 30  
 Asn Ala Gln Arg Asn Ala Leu Asp Asn Glu Ile Thr Xaa Ala Thr Asn  
 35 40 45  
 Val Glu Val Leu Ile Gln Leu Lys Pro Lys Arg Asn Asn  
 50 55 60

(2) INFORMATION FOR SEQ ID NO:362:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 241 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single



# EP 0 841 394 A2

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:362:

10	Val	Ile	Glu	Met	Ala	Lys	Lys	Ala	Pro	Asp	Val	Gly	Asp	Tyr	Lys	Tyr
	1				5				10					15		
	Gly	Phe	Pro	Ala	Asp	Asp	Val	Ser	Ile	Phe	Arg	Ser	Glu	Arg	Gly	Leu
				20				25					30			
15	Thr	Glu	Asn	Ile	Val	Arg	Glu	Ile	Ser	Asn	Met	Lys	Asn	Glu	Pro	Glu
			35				40					45				
	Trp	Met	Leu	Asp	Phe	Arg	Leu	Lys	Ser	Leu	Lys	Leu	Phe	Tyr	Lys	Met
	50						55				60					
20	Pro	Met	Pro	Gln	Trp	Gly	Gly	Asp	Leu	Ser	Glu	Leu	Asn	Phe	Asp	Asp
	65				70				75				80			
	Ile	Thr	Tyr	Tyr	Val	Lys	Pro	Ser	Glu	Gln	Ala	Glu	Arg	Ser	Trp	Asp
				85				90				95				
25	Glu	Val	Pro	Glu	Glu	Ile	Lys	Arg	Thr	Phe	Asp	Lys	Leu	Gly	Ile	Pro
			100				105				110					
	Glu	Ala	Glu	Gln	Lys	Tyr	Leu	Ala	Gly	Val	Ser	Ala	Gln	Tyr	Glu	Ser
		115					120				125					
30	Glu	Val	Val	Tyr	His	Asn	Met	Glu	Lys	Glu	Leu	Glu	Glu	Lys	Gly	Ile
	130					135				140						
	Ile	Phe	Lys	Asp	Thr	Asp	Ser	Ala	Leu	Gln	Glu	Asn	Glu	Glu	Leu	Phe
	145				150				155				160			
35	Lys	Lys	Tyr	Phe	Ala	Ser	Val	Val	Pro	Ala	Ala	Asp	Asn	Lys	Phe	Ala
				165				170				175				
	Ala	Leu	Asn	Ser	Ala	Val	Trp	Ser	Gly	Xaa	Ser	Leu	Ile	Tyr	Val	Pro
		180					185				190					
40	Lys	Asn	Ile	Lys	Leu	Asp	Thr	Pro	Leu	Gln	Ala	Tyr	Phe	Arg	Ile	Asn
		195				200				205						
	Ser	Glu	Asn	Met	Gly	Gln	Phe	Glu	Arg	Thr	Leu	Ile	Ile	Ala	Asp	Glu
	210					215				220						
45	Gly	Ala	Ser	Val	His	Tyr	Val	Glu	Gly	Trp	Tyr	Cys	Thr	Ser	Leu	Tyr
	225				230				235			240				
	Asn															

(2) INFORMATION FOR SEQ ID NO:363:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 275 amino acids

(B) TYPE: amino acid

EP 0 841 394 A2

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:363:

10

```

Val Lys Gln His Lys Ala Trp Val Asn Ile Met Tyr Gly Cys Asp Lys
 1             5             10             15
Phe Cys Thr Tyr Cys Ile Val Pro Xaa Thr Arg Ser Lys Asp Arg Thr
                20             25             30
15 Arg Arg Pro Glu Asp Ile Ile Asp Glu Val Arg Glu Leu Ala Arg Glu
    35             40             45
Gly Tyr Lys Glu Ile Thr Leu Leu Gly His Asn Val Asn Ser Tyr Gly
    50             55             60
20 Lys Asp Leu Gln Asp Ile Glu Tyr Asp Leu Gly Asp Leu Leu Gln Ala
    65             70             75             80
Ile Ser Lys Ile Ala Ile Pro Arg Val Arg Phe Thr Thr Ser His Pro
                85             90             95
25 Trp Asp Phe Thr Asp His Met Ile Asp Val Ile Ser Glu Gly Gly Asn
    100            105            110
Ile Val Pro His Ile His Leu Pro Val Gln Ser Gly Asn Asn Ala Val
    115            120            125
30 Leu Lys Ile Met Gly Arg Lys Tyr Thr Arg Glu Ser Tyr Leu Asp Leu
    130            135            140
Val Lys Arg Ile Lys Asp Arg Leu Pro Asn Val Ala Leu Thr Thr Asp
    145            150            155            160
35 Ile Ile Val Gly Tyr Pro Asn Glu Ser Glu Glu Gln Phe Glu Glu Thr
    165            170            175
Leu Thr Leu Tyr Asp Glu Val Gly Phe Glu His Ala Tyr Thr Tyr Leu
    180            185            190
40 Tyr Ser Gln Arg Asp Gly Thr Pro Ala Ala Lys Met Lys Asp Asn Val
    195            200            205
Pro Leu Asn Val Lys Lys Glu Arg Leu Gln Arg Leu Asn Lys Lys Val
    210            215            220
45 Gly His Tyr Ser Gln Ile Ala Met Ser Lys Tyr Glu Gly Gln Thr Val
    225            230            235            240
Thr Val Leu Cys Glu Gly Ser Ser Lys Lys Asp Asp Gln Val Leu Ala
    245            250            255
50 Gly Tyr Thr Asp Lys Asn Lys Leu Val Asn Phe Lys Ala Pro Lys Glu
    260            265            270
Met Ile Gly
55      275

```

EP 0 841 394 A2

(2) INFORMATION FOR SEQ ID NO:364:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 62 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:364:

Val	Thr	Gln	Tyr	Leu	Val	Thr	Thr	Phe	Lys	Asp	Ser	Thr	Gly	Arg	Pro
1				5					10					15	
His	Glu	His	Ile	Thr	Val	Ala	Arg	Asp	Asn	Gln	Thr	Phe	Thr	Val	Ile
			20					25						30	
Glu	Ala	Glu	Ser	Lys	Glu	Glu	Ala	Glu	Arg	Lys	Tyr	Glu	Ala	Gln	Val
			35					40						45	
Lys	Ile	Arg	Arg	Asp	Gly	Asp	Ala	Lys	Glu	Asn	Xaa	Asn	Asp		
			50					55						60	

(2) INFORMATION FOR SEQ ID NO:365:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:365:

Val	Tyr	Lys	Lys	Ala	Gln	Ala	Phe	Asp	Glu	Ile	Leu	Glu	Gly	Met	Thr
1				5					10					15	
Asn	Ala	Ile	Gln	His	Ser	Val	Lys	Xaa	Gly	Ile	Glu	Leu	Asp	Glu	Ala
			20					25						30	
Val	Gly	Ile	Met	Ala	Gly	Gln	Val	Val	Tyr	Lys	Tyr	Glu	Glu	Glu	
			35					40						45	

(2) INFORMATION FOR SEQ ID NO:366:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 105 amino acids

# EP 0 841 394 A2

(B) TYPE: amino acid  
(C) STRANDEDNESS: single.  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:366:

```

Val Ile Asp Met Lys Phe Lys Lys Val Leu Val Ala Thr Ala Met Val
 1             5             10             15
Gly Val Leu Ala Thr Gly Val Val Gly Tyr Gly Asn Gln Ala Asp Ala
15             20             25             30
Lys Val Tyr Ser Gln Asn Gly Leu Val Leu His Asp Asp Ala Asn Phe
             35             40             45
Leu Glu His Glu Leu Ser Tyr Ile Asp Val Leu Leu Asp Lys Asn Ala
20             50             55             60
Asp Gln Ala Thr Lys Asp Asn Leu Arg Ser Tyr Phe Ala Asp Lys Gly
65             70             75             80
Leu His Ser Ile Lys Asp Ile Ile Asn Lys Ala Lys Gln Asp Gly Phe
25             85             90             95
Asp Val Ser Lys Tyr Glu His Val Lys
             100             105

```

(2) INFORMATION FOR SEQ ID NO:367:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 112 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:367:

```

Val Thr Asn Lys Glu Tyr Glu Ile Glu Pro Gly Lys Arg Glu Cys Glu
 1             5             10             15
Met Met Lys Arg Leu Asn Lys Leu Val Leu Gly Ile Ser Phe Leu Phe
50             20             25             30
Leu Val Ile Ser Ile Thr Ala Gly Cys Gly Ile Gly Lys Glu Ala Glu
             35             40             45
Val Lys Lys Ser Phe Glu Lys Thr Leu Ser Met Tyr Pro Ile Lys Asn
55             50             55             60

```

EP 0 841 394 A2

Leu Glu Asp Leu Tyr Asp Lys Glu Gly Tyr Arg Asp Asp Gln Phe Asp  
 65 70 75 80  
 Lys Asn Asp Lys Gly Thr Trp Ile Ile Asn Ser Glu Met Val Ile Gln  
 85 90 95  
 Pro Asn Asn Glu Asp Met Val Ala Lys Gly Met Val Leu Tyr Ile Glu  
 100 105 110

(2) INFORMATION FOR SEQ ID NO:368:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 104 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:368:

Val Ile Asn Ile Pro Lys Met Leu Lys Tyr Asp Gln Ile Leu Val Tyr  
 1 5 10 15  
 Ser Asn Pro Pro Ile Leu Pro Leu Ile Pro Asp Val Leu His Arg Leu  
 20 25 30  
 Leu Lys Lys Lys Tyr Ser Phe Val Val Tyr Asp Ile Ala Pro Asp Asn  
 35 40 45  
 Ala Ile Lys Thr Gly Ala Thr Arg Pro Gly Ser Met Ile Asp Lys Leu  
 50 55 60  
 Met Arg Tyr Ile Asn Arg His Val Tyr Lys Asn Ala Glu Asn Val Ile  
 65 70 75 80  
 Val Leu Gly Tyr Gly Asn Glu Lys Leu Leu Thr Lys Ser Ser Asn Phe  
 85 90 95  
 Leu Lys Met Leu Thr Ile Ser Met  
 100

(2) INFORMATION FOR SEQ ID NO:369:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 88 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:369:

```

5   Val Lys Tyr Ile Glu Ala Pro Tyr Glu Pro His Lys Phe Val Lys Met
    1             5             10             15
    Val Lys Asp Lys Glu Leu Ala Asp Glu Lys Glu Gly Gly Leu Arg Xaa
      20             25             30
10   Thr Ala Cys Phe Glu Met Arg Leu Asp Ile Val Ala Lys Ala Ala Val
    35             40             45
    Glu His Gly Tyr Asp Tyr Phe Gly Ser Ala Ile Thr Leu Ser Pro Lys
    50             55             60
15   Lys Asn Ala Gln Leu Ile Asn Glu Leu Gly Met Asp Cys Pro Lys Asn
    65             70             75             80
    Ile Arg Cys Glu Leu Ile Cys Lys
      85

```

(2) INFORMATION FOR SEQ ID NO:370:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 119 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:370:

```

35   Val Thr Glu Ile Lys Tyr Lys Val Ile Thr Lys Asp Ala Phe Ala Leu
    1             5             10             15
    Pro Tyr Thr Ile Ile Lys Ala Lys Asn Gln Pro Thr Lys Gly Val Ile
      20             25             30
40   Val Tyr Ile His Gly Gly Gly Leu Met Phe Gly Lys Ala Asn Asp Leu
    35             40             45
    Ser Pro Gln Tyr Ile Asp Ile Leu Thr Glu His Tyr Asp Leu Ile Gln
    50             55             60
45   Leu Ser Tyr Arg Leu Leu Pro Glu Val Ser Leu Asp Cys Ile Ile Glu
    65             70             75             80
    Asp Val Tyr Ala Ser Phe Asp Ala Ile Gln Ser Gln Tyr Ser Asn Cys
      85             90             95
50   Pro Ile Phe Thr Phe Gly Arg Ser Ser Gly Ala Tyr Leu Ser Leu Ile
    100             105             110
    Asn Cys Thr Arg Gln Arg Tyr
      115

```

EP 0 841 394 A2

(2) INFORMATION FOR SEQ ID NO:371:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 166 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:371:

Val	Ile	Asp	Val	Tyr	Asn	Gln	Ile	Lys	Asn	Asn	Leu	Glu	Ser	Met	Thr
1				5				10						15	
Pro	Glu	Thr	Ala	Thr	Ile	Gly	Arg	Leu	Val	Asp	Thr	Lys	Thr	Leu	Phe
			20					25						30	
Leu	Glu	Tyr	Arg	Lys	Lys	Leu	Gln	Asp	Val	Tyr	Thr	Asp	Val	Glu	Asp
			35					40						45	
Val	Lys	Ile	Ala	Ile	Ser	Asp	Arg	Phe	Lys	Leu	Leu	Gln	Ser	Gln	Tyr
			50					55						60	
Thr	Asp	Glu	Lys	Tyr	Lys	Glu	Ala	Leu	Glu	Ile	Ile	Ala	Thr	Lys	Phe
			65					70						75	
Gly	Leu	Thr	Val	Asn	Glu	Asp	Leu	Gln	Leu	Val	Gly	Glu	Pro	Asn	Val
			85					90						95	
Val	Lys	Ser	Ala	Ile	Glu	Ala	Ala	Arg	Glu	Ser	Thr	Lys	Glu	Gln	Leu
			100					105						110	
Arg	Asp	Tyr	Xaa	Lys	Thr	Ser	Asp	Tyr	Xaa	Thr	Asp	Xaa	Asp	Gly	Ile
			115					120						125	
Val	Glu	Arg	Xaa	Asp	Thr	Ala	Glu	Ala	Glu	Arg	Thr	Thr	Xaa	Xaa	Gly
			130					135						140	
Glu	Ile	Lys	Asp	Lys	Xaa	Thr	Val	Xaa	Arg	Ile	Ser	Lys	Arg	Ile	Gly
			145					150						155	
Arg	Thr	Lys	Thr	Ile	Tyr										160
															165

(2) INFORMATION FOR SEQ ID NO:372:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 303 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:372:

5	Val	Leu	Ala	Asp	Ile	Thr	Val	Asn	Ala	Met	Lys	Gly	Ile	Tyr	Leu	Arg
	1				5					10					15	
	Tyr	Asp	Glu	Asn	Gly	Ala	Ile	Thr	Ser	His	Thr	Ile	Asp	Lys	Asp	Gly
				20					25					30		
10	Val	Lys	Ile	Ser	Gly	Asp	Lys	Val	Asp	Ile	Thr	Ala	Asn	Arg	Glu	Phe
		35						40					45			
	Asn	Val	Phe	Ala	Asn	Asn	Ile	Asn	Asn	Lys	Val	Gly	Lys	Asn	Asp	Ile
		50					55					60				
15	Val	Asn	Ser	Leu	Asn	Leu	Ser	Asn	Glu	Gly	Leu	Asp	Ile	Asn	Val	Asn
	65				70					75					80	
	Arg	Ile	Gly	Ile	Lys	Gly	Gly	Asn	Ala	Asn	Arg	Tyr	Val	Gln	Val	Gln
					85					90				95		
20	Asn	Asp	Phe	Ile	Glu	Leu	Gly	Gly	Ile	Val	Gln	Arg	Thr	Trp	Lys	Gly
			100						105					110		
	Lys	Arg	Ser	Thr	Asp	Asp	Ile	Phe	Thr	Arg	Leu	Lys	Asp	Gly	His	Leu
			115				120						125			
25	Arg	Phe	Arg	Asn	Asn	Thr	Ala	Gly	Gly	Ser	Leu	Tyr	Met	Ser	His	Phe
		130				135						140				
	Gly	Ile	Ser	Thr	Tyr	Ile	Asp	Gly	Glu	Gly	Glu	Asp	Gly	Gly	Ser	Ser
	145					150				155					160	
30	Gly	Thr	Ile	Gln	Trp	Trp	Asp	Lys	Thr	Tyr	Ser	Asp	Ser	Gly	Met	Asn
				165						170				175		
	Gly	Ile	Thr	Ile	Asn	Ser	Tyr	Gly	Gly	Val	Val	Ala	Leu	Thr	Ser	Asp
				180					185					190		
35	Tyr	Asn	Arg	Ile	Ile	Ile	Asp	Ser	Tyr	Ala	Ser	Ala	Asn	Ile	Glu	Ser
		195					200						205			
	Arg	Glu	Ala	Pro	Ile	Tyr	Leu	Ser	Pro	Asn	Thr	Gln	Lys	Leu	Asn	Leu
		210					215					220				
40	Xaa	Leu	Xaa	Arg	Phe	Ala	Phe	Thr	Leu	Ser	Asn	Ala	Asp	Arg	Xaa	Ile
	225				230						235				240	
	Thr	Lys	Leu	Ala	Val	Ile	Ser	Cys	Trp	Val	Gln	Asp	Xaa	Xaa	Tyr	Lys
				245					250					255		
45	Xaa	Gly	Ala	Gly	Leu	Arg	Phe	Ser	Lys	Arg	Thr	Asn	Lys	Gly	Leu	Val
			260						265					270		
	Gln	Val	Val	Asn	Gly	Asp	Tyr	Ala	Thr	Gly	Gly	Asp	Thr	Thr	Ile	Glu
			275					280					285			
50	Ser	Gly	Met	Ala	Lys	Phe	Asn	Leu	Val	Xaa	Arg	Lys	Arg	Trp	Lys	
		290					295					300				

(2) INFORMATION FOR SEQ ID NO:373:



(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 95 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:373:

```

Val Asn Pro Thr His Val Lys Tyr Ala Ala Glu Arg Leu Ala Asp Ser
 1           5           10          15
Glu Val Leu Val Cys Thr Val Ile Gly Phe Pro Leu Gly Ala Ser Thr
      20          25          30
Thr Ala Thr Lys Ala Phe Glu Thr Glu Asp Ala Ile Gln Asn Gly Ala
 20      35          40          45
Asp Glu Ile Asp Met Val Ile Asn Ile Gly Ala Leu Lys Asp Gly Arg
      50          55          60
Phe Asp Asp Val Gln Gln Asp Ile Glu Ala Val Val Lys Ala Ala Lys
 25      65          70          75          80
Gly His Thr Val Lys Val Ile Ile Glu Thr Val Leu Leu Glu Pro
      85          90          95

```

(2) INFORMATION FOR SEQ ID NO:374:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:374:

```

Val Lys Ile Tyr Glu Asp Thr Asn Ile Asp Thr Leu Xaa Leu Leu Asn
 1           5           10          15
Glu Ala Xaa Ile Phe Lys Xaa Thr Leu Phe Trp Cys Xaa Xaa Ser Asn
      20          25          30

```

(2) INFORMATION FOR SEQ ID NO:375:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 55 amino acids

EP 0 841 394 A2

(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:375:

Val	Ala	His	Asp	Asn	His	Ile	Gly	Thr	Tyr	Cys	Ile	Val	Met	Ser	Gly
1				5				10				15			
Arg	Gly	Pro	Ser	Asp	Lys	Glu	Val	Asp	His	Ile	Ser	Asn	Pro	Val	Arg
			20					25				30			
Thr	Ile	Lys	Ser	Gln	His	Pro	Gln	Leu	Lys	Ile	Cys	Ala	Cys	Leu	Gly
			35					40				45			
Leu	Thr	Asp	Cys	Arg	Pro	Ser									
			50					55							

(2) INFORMATION FOR SEQ ID NO:376:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 66 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:376:

Val	Ile	Asp	Asp	Leu	Val	Ser	Asp	His	Ile	Ala	Thr	Gly	Gln	Leu	Leu
1				5				10				15			
Val	Lys	Met	Ser	Asp	Leu	Thr	Ser	Ser	Tyr	Glu	Pro	Pro	Ile	Glu	Ala
			20					25				30			
Cys	Gly	Thr	Trp	Arg	Leu	Val	Tyr	Gln	Arg	Leu	Lys	Ala	Leu	Glu	Val
			35					40				45			
Leu	Thr	His	Glu	His	Val	His	Leu	Glu	Asn	His	Val	Leu	Phe	Lys	Lys
			50					55				60			
Val	Ser														

(2) INFORMATION FOR SEQ ID NO:377:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 54 amino acids

(B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:377:

```

Val Leu Met Ile Phe Gly Pro Glu Gly Gly Leu Ser Glu Ile Glu Ile
 1             5             10             15
Ser Leu Phe Ser Asn Thr Ser Thr Val Val Gly Leu Gly Pro Arg Ile
15             20             25             30
Leu Arg Ala Glu Thr Ala Pro Leu Tyr Ala Leu Ser Ala Ile Ser Tyr
35             40             45
Glu Lys Glu Leu Met Gly
20             50
  
```

(2) INFORMATION FOR SEQ ID NO:378:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 134 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:378:

```

Val Ser Ser Ile Trp Lys Glu Lys Glu Phe Glu Pro Ser Asp Ile Val
40 1             5             10             15
Asp Ala Tyr Leu Val Ile Ala Ala Thr Asn Glu Pro Arg Val Asn Glu
20             25             30
Ala Val Lys Gln Ala Leu Pro Glu His Ala Leu Phe Asn Asn Val Gly
45 35             40             45
Asp Ala Ser Asn Ala Asn Val Val Phe Pro Ser Ala Leu His Arg Asp
50 50             55             60
Lys Leu Thr Ile Ser Val Ser Thr Asp Gly Ala Ser Pro Lys Leu Thr
50 65             70             75             80
Lys Ser Ile Met Ala Glu Leu Glu Ala Leu Tyr Pro Pro Ser Tyr Ser
85             90             95
Ser Tyr Ile Asp Phe Leu Tyr Thr Cys Arg Gln Lys Ile Lys Val Leu
55 100             105             110
  
```

EP 0 841 394 A2

Asp Ile Thr Tyr Ile Arg Lys Ala Thr Val Thr Val Thr Asn Cys Val  
 115 120 125  
 5 Thr Arg Ile Phe Lys Ser  
 130

(2) INFORMATION FOR SEQ ID NO:379:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 244 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:379:

Val Val Asn His Leu Cys His Gly Glu Asp Gln Asp Val Xaa Ala His  
 1 5 10 15  
 Ala Ile Thr Asp Leu Glu Asn Lys Gly Ile Ala Leu His Thr Asn Val  
 20 25 30  
 Glu Thr Thr Glu Leu Ser Ser Asp Asp His His Thr Thr Val Arg Xaa  
 35 40 45  
 Asn Val Gly Asn Xaa Glu Ala Asp Ala Val Leu Leu Ala Ile Gly Arg  
 50 55 60  
 Lys Pro Asn Thr Ala Leu Ala Leu Glu Asn Thr Asp Ile Glu Leu Gly  
 65 70 75 80  
 Asp Arg Gly Glu Ile Lys Val Asn Ala Gln Leu Gln Thr Ser Val Pro  
 85 90 95  
 His Ile Tyr Ala Ala Gly Asp Val Lys Gly Gly Leu Gln Phe Thr Tyr  
 100 105 110  
 Ile Ser Leu Asp Asp Tyr Arg Ile Ile Lys Ser Ala Leu Tyr Gly Asn  
 115 120 125  
 Gln Ser Arg Thr Thr Asp Asn Arg Gly Ser Val Pro Tyr Thr Val Phe  
 130 135 140  
 Ile Asp Pro Pro Leu Ser Arg Val Gly Leu Thr Ser Lys Glu Ala Ala  
 145 150 155 160  
 Ala Gln His Tyr Asp Tyr Thr Glu His Gln Leu Leu Val Ser Ala Ile  
 165 170 175  
 Pro Arg His Lys Ile Asn Asn Asp Pro Arg Gly Leu Phe Lys Val Val  
 180 185 190  
 Ile Asn Asn Glu Asn Asn Met Ile Leu Gly Ala Thr Leu Tyr Gly Lys  
 195 200 205

EP 0 841 394 A2

Gln Ser Glu Glu Leu Ile Asn Ile Ile Lys Leu Ala Ile Asp Gln Asn  
 210 215 220  
 5 Ile Pro Tyr Thr Val Leu Arg Asp Ser Ile Tyr Ser His Pro Thr Met  
 225 230 235 240  
 Gly Arg Ile Ile

(2) INFORMATION FOR SEQ ID NO:380:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:380:

Val Val Glu Trp Ala Leu Tyr Ile Ala Lys Asn Lys Ile Ala Ile Asp  
 1 5 10 15  
 Val Pro Gly Ser Gly Met Gly Ala Gln Cys Trp Glu Phe Thr  
 20 25 30

(2) INFORMATION FOR SEQ ID NO:381:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 93 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:381:

Val Leu Glu Leu Arg Glu Glu His Lys Gln His His Asn Glu Leu Arg  
 1 5 10 15  
 Glu Ser His Lys Glu Leu Lys Asp Lys Gln Asp Lys Val Val Asp Glu  
 20 25 30  
 Asn Leu Glu Gln Thr Lys Ile Leu Asn Arg Ile Glu Glu Arg Tyr Xaa  
 35 40 45  
 Thr Gln Val Xaa Val Xaa Gln Lys Asn Glu Glu Lys Thr Leu Ala Gln  
 50 55 60

EP 0 841 394 A2

Asn Lys Trp Leu Val Gly Ala Ile Trp Ala Leu Val Thr Ile Val Met  
 65 70 75 80  
 5 Ile Ala Val Ile Thr Ala Ser Ile Xaa Ala Leu Leu Pro  
 85 90

(2) INFORMATION FOR SEQ ID NO:382:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 174 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:382:

Val Asn Ser Thr Leu Ile Arg Ile Gly Asp Ile Ile Gln Ser Ser Thr  
 1 5 10 15  
 Asn Thr Asn Pro Ile Ile Met Gly Ile Ile Leu Gly Gly Ile Ile Thr  
 20 25 30  
 Val Val Gly Thr Ala Pro Leu Ser Ser Met Ala Leu Thr Ala Leu Leu  
 35 40 45  
 Gly Leu Thr Gly Val Pro Met Ala Ile Gly Ala Met Ala Ala Phe Ser  
 50 55 60  
 Ser Ala Phe Met Asn Gly Thr Leu Phe His Arg Leu Lys Leu Gly Asp  
 65 70 75 80  
 Arg Lys Ser Thr Ile Ala Val Ser Ile Glu Pro Leu Ser Gln Ala Asp  
 85 90 95  
 Ile Val Ser Ala Asn Pro Ile Pro Ile Tyr Ile Thr Asn Phe Phe Gly  
 100 105 110  
 Gly Ala Ile Ala Gly Leu Ile Ile Ala Met Ser Gly Leu Ile Asn Asp  
 115 120 125  
 Ala Thr Gly Thr Ala Thr Pro Ile Xaa Gly Phe Leu Val Met Xaa Gly  
 130 135 140  
 Phe Asn His Pro Met Thr Ile Val Ile Tyr Gly Val Val Met Ala Ile  
 145 150 155 160  
 Val Gly Cys Ala Cys Arg Val Ile Leu Gly Ser Ile Arg Ile  
 165 170

(2) INFORMATION FOR SEQ ID NO:383:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 76 amino acids

EP 0 841 394 A2

(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:383:

Val	Gly	Lys	Thr	Gly	Phe	Ala	Leu	Asn	Met	Met	Leu	Asn	Ile	Ala	Arg
1				5				10					15		
Asn	Gly	Tyr	Lys	Thr	Ser	Phe	Phe	Ser	Leu	Glu	Thr	Thr	Gly	Thr	Ser
			20					25					30		
Val	Leu	Lys	Arg	Met	Leu	Ser	Thr	Ile	Thr	Gly	Ile	Glu	Leu	Thr	Lys
			35				40					45			
Ile	Lys	Glu	Ile	Arg	Asn	Leu	Thr	Pro	Asp	Asp	Leu	Thr	Lys	Leu	Thr
		50				55				60					
Asn	Ala	Met	Gly	Ser	Lys	Ser	Leu	Lys	Leu	Gly	Ile				
65					70					75					

(2) INFORMATION FOR SEQ ID NO:384:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 130 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:384:

Val	Pro	Asn	Glu	Thr	Glu	Ala	Glu	Leu	Leu	Ser	Gly	Ile	Lys	Val	Thr
1				5				10					15		
Asn	Glu	Gln	Ser	Met	Lys	Asp	Asn	Ala	Asn	Tyr	Phe	Leu	Ser	Leu	Gly
			20					25					30		
Ile	Lys	Thr	Val	Leu	Ile	Thr	Leu	Gly	Lys	Gln	Gly	Thr	Tyr	Phe	Ala
			35				40					45			
Thr	Lys	Asn	Gln	Ser	Gln	His	Ile	Glu	Ala	Tyr	Lys	Val	Asn	Ala	Ile
		50				55				60					
Asp	Thr	Thr	Ala	Ala	Gly	Asp	Thr	Phe	Ile	Gly	Ala	Phe	Val	Ser	Arg
65					70					75				80	
Leu	Asn	Lys	Ser	Gln	Asp	Asn	Leu	Ala	Asp	Ala	Ile	Asp	Phe	Gly	Asn
				85				90						95	

# EP 0 841 394 A2

Lys Ala Ser Ser Leu Thr Val Gln Lys Thr Arg Arg Ala Ser Ile Tyr  
 100 105 110  
 Ser Ser Thr Arg Arg Ser Lys Ser Ser Leu Asn Glu Ser Asn Thr Ala  
 115 120 125  
 Met Ile  
 130

## (2) INFORMATION FOR SEQ ID NO:385:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 172 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Protein

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:385:

Val Tyr Lys Met Arg Lys Ile Gly Tyr Ala Arg Val Ala Tyr Pro Asp  
 1 5 10 15  
 Gln Asn Leu Asp Thr Gln Leu Thr Lys Leu Leu Ile Asn Gly Cys Asp  
 20 25 30  
 Leu Val Tyr Ser Glu Gln Val Asn Val Tyr Tyr Lys Glu Gln Leu Glu  
 35 40 45  
 Leu Glu His Cys Leu Asp Glu Leu Lys Thr Asp Asp Thr Leu Val Ile  
 50 55 60  
 Glu Lys Leu Lys Val Leu Gly Phe Thr Pro Lys Lys Leu Met Glu Phe  
 65 70 75 80  
 Phe Glu Ser Arg Ile Leu Pro Tyr Asp Ile His Leu Glu Val Leu Asp  
 85 90 95  
 Leu Gly Ile Asn Thr Asn Ser Glu Glu Gly Gln Ser Phe Ile Glu Val  
 100 105 110  
 Phe Lys Met Leu Ala Asp Ser Glu Asn Ile Leu Leu Lys Glu Arg Thr  
 115 120 125  
 Thr Asn Gly Leu Glu Ser Ala Lys Glu Arg Gly Arg Tyr Gly Gly Arg  
 130 135 140  
 Pro Gln Leu Ser Glu Asp Lys Arg Lys Tyr Ile Lys Gln Leu Phe Ala  
 145 150 155 160  
 Ser Arg Met Tyr Thr Pro Asn Glu Ile Ser Lys Trp  
 165 170

## (2) INFORMATION FOR SEQ ID NO:386:



# EP 0 841 394 A2

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 300 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:386:

15	Val	Leu	Ile	Met	Ile	Asn	Glu	Met	Leu	Asn	Pro	Lys	Gln	Gln	Glu	Val
	1				5					10					15	
	Trp	Thr	Cys	Phe	Ile	Asn	Asp	Lys	Pro	Lys	Val	Leu	Ile	Ala	Ser	Gly
				20					25					30		
20	Ala	Lys	Arg	Ala	Gly	Lys	Thr	Tyr	Val	Phe	Ile	Leu	Leu	Phe	Leu	Met
			35					40					45			
	His	Ile	Ala	Thr	Tyr	Lys	Asp	Lys	Gly	Leu	Asn	Phe	Ile	Ile	Gly	Gly
		50				55					60					
25	Ala	Thr	Gln	Ala	Ser	Ile	Arg	Arg	Asn	Ile	Leu	Asp	Asp	Met	Glu	Leu
	65					70				75					80	
	Ile	Leu	Gly	Arg	Glu	Leu	Thr	Leu	Asp	Lys	Ser	Asn	Ala	Val	Lys	Ile
				85					90					95		
30	Phe	Gly	Asn	Lys	Val	Tyr	Val	Phe	Asp	Gly	Gln	Asn	Ser	Asp	Ala	Trp
			100						105					110		
	Lys	Lys	Ala	Arg	Gly	Phe	Thr	Ser	Ala	Gly	Ala	Phe	Leu	Asn	Glu	Gly
			115				120						125			
35	Thr	Ala	Leu	His	Asn	Met	Phe	Ile	Lys	Glu	Val	Phe	Ser	Arg	Cys	Ser
		130				135					140					
	Tyr	Lys	Gly	Ala	Arg	Ile	Leu	Ile	Asp	Thr	Asn	Pro	Glu	Asn	Pro	Met
	145				150					155					160	
40	His	Pro	Val	Lys	Lys	Asp	Tyr	Ile	Asp	Lys	Ser	Gly	Gln	Arg	Leu	Ser
				165					170					175		
	Asn	Gly	Arg	Leu	Asn	Ile	Lys	Ala	Phe	Gln	Phe	Thr	Leu	Phe	Asp	Asn
			180					185					190			
45	Thr	Phe	Leu	Asp	Glu	Glu	Tyr	Ile	Glu	Ser	Ile	Ile	Ala	Ser	Thr	Pro
		195					200						205			
	Thr	Gly	Met	Phe	Thr	Asp	Arg	Asp	Ile	Tyr	Gly	Lys	Trp	Val	Ser	Ala
		210					215					220				
50	Glu	Gly	Val	Val	Tyr	Lys	Asp	Phe	Lys	Glu	Lys	Val	His	Tyr	Ile	Thr
	225				230					235				240		
	Glu	Glu	Glu	Phe	Lys	Thr	Lys	Gln	Ile	Lys	Arg	Lys	Tyr	Ala	Gly	Val
				245				250					255			
55	Asp	Trp	Gly	Tyr	Glu	His	Tyr	Gly	Ser	Ile	Met	Val	Val	Ala	Glu	Asp
				260				265					270			

EP 0 841 394 A2

Phe Asp Gly Asn Lys Tyr Val Ile Glu Glu His Ala His Arg His Lys.  
 275 280 285  
 5 Glu Ile Asp Asp Trp Val Ala Ile Ala Lys Glu Leu  
 290 295 300

(2) INFORMATION FOR SEQ ID NO:387:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 51 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:387:

Val Asn Leu Ile Arg Ser His Ala Cys Gly Leu Gly Glu Pro Phe Pro  
 1 5 10 15  
 Lys Glu Val Ala Leu Val Met Met Ile Xaa Arg Leu Asn Thr Leu Leu  
 20 25 30  
 Lys Gly His Ser Gly Ala Thr Leu Val Ile Ser Glu Thr Ile Thr Ile  
 35 40 45  
 Phe Tyr Lys  
 50

(2) INFORMATION FOR SEQ ID NO:388:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 199 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:388:

Val Ile His Val Leu Asp Phe Asn Asp Lys Ile Ile Asp Phe Leu Ser  
 1 5 10 15  
 Thr Asp Asp Pro Ser Leu Val Arg Ala Ile His Lys Arg Asn Val Asn  
 20 25 30  
 Asp Asn Ser Glu Met Leu Glu Leu Leu Ile Ser Ser Glu Arg Ala Glu  
 35 40 45

EP 0 841 394 A2

Xaa Phe Arg Glu Arg His Arg Val Ile Ile Arg Asp Ser Asn Lys Gln  
50 55 60  
5 Trp Arg Glu Phe Ile Ile Asn Trp Val Gln Asp Thr Met Asp Gly Tyr  
65 70 75 80  
Thr Glu Ile Glu Cys Ile Ala Ser Tyr Leu Ala Asp Ile Thr Thr Ala  
85 90 95  
10 Lys Pro Tyr Ala Pro Gly Lys Phe Glu Lys Lys Thr Thr Ser Glu Ala  
100 105 110  
Leu Lys Asp Val Leu Ser Asp Thr Gly Trp Glu Val Ser Glu Xaa Thr  
115 120 125  
15 Glu Tyr Asp Gly Leu Arg Thr Thr Ser Trp Thr Ser Tyr Xaa Thr Arg  
130 135 140  
Tyr Glu Val Leu Lys Gln Leu Cys Thr Thr Tyr Lys Met Ala Leu Asp  
145 150 155 160  
20 Phe Tyr Ile Glu Leu Ser Ser Asn Thr Val Lys Gly Arg Tyr Val Val  
165 170 175  
Leu Asn Lys Lys Asn Ser Leu Phe Gln Gly Lys Glu Ile Glu Tyr Gly  
180 185 190  
25 Lys Asp Leu Ala Trp Val Asn  
195

(2) INFORMATION FOR SEQ ID NO:389:

30 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 156 amino acids  
(B) TYPE: amino acid  
35 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:389:

Val Leu Val Gln Gly Gln Val Cys Ser His Leu Ser Thr Xaa Leu Gly  
1 5 10 15  
45 Leu Ile Ile Ile Asp Glu Glu His Glu Ser Thr Tyr Lys Gln Glu Asp  
20 25 30  
Tyr Pro Arg Tyr His Ala Arg Glu Ile Ala Gln Trp Arg Ser Glu Tyr  
35 40 45  
50 His His Cys Pro Val Ile Leu Gly Ser Ala Thr Pro Cys Leu Glu Ser  
50 55 60  
Tyr Ala Arg Ala Glu Lys Asp Val Tyr His Leu Leu Ser Leu Pro Asn  
65 70 75 80

EP 0 841 394 A2

Arg Val Asn Gln Gln Ala Leu Pro Glu Ile Asp Ile Val Asp Met Arg  
 85 90 95  
 5 Glu Glu Leu Ser Glu Gly Asn Arg Ser Met Phe Ser Lys Asp Leu Arg  
 100 105 110  
 Glu Ala Ile Gln Leu Arg Xaa Arg Ser Thr Gly Thr Ser Cys Phe Ile  
 115 120 125  
 10 Phe Lys Ser Asp Val Val Met His Arg Leu Cys Tyr Val Gly Ile Val  
 130 135 140  
 Asp Met Tyr Arg Asn Val Gln Xaa Val Ile Phe His  
 145 150 155

(2) INFORMATION FOR SEQ ID NO:390:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 88 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:390:

Val Lys Tyr Leu Asn Asp Asp Ile Ala Lys Gly Ser Ile Phe Asp Tyr  
 1 5 10 15  
 Leu Glu Ser Asn Met Lys Leu Arg Ile Gly Phe Ser Asp Ile Phe Phe  
 20 25 30  
 35 Asn Val Asp Lys Leu Thr Ser Ser Glu Ala Ser Leu Leu Gln Leu Ser  
 35 40 45  
 Thr Gly Glu Pro Cys Leu Arg Tyr His Gln Thr Phe Tyr Thr Met Thr  
 50 55 60  
 40 Gly Lys Pro Phe Asp Ser Ser Asp Ile Val Phe His Tyr Xaa His Ala  
 65 70 75 80  
 Gln Phe Tyr Ile Pro Ser Lys Lys  
 85

(2) INFORMATION FOR SEQ ID NO:391:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 110 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

EP 0 841 394 A2

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:391:

```

Val Arg His Leu His Ile Pro Leu Gln Ser Gly Ser Asp Thr Val Leu
 1           5           10           15
Lys Arg Met Arg Arg Lys Tyr Thr Met Asp Arg Phe Ser Glu Arg Leu
10          20          25          30
Thr Lys Leu His Lys Ala Leu Pro Asp Leu Ala Val Thr Ser Asp Val
          35          40          45
Ile Val Gly Phe Pro Gly Glu Thr Glu Ala Glu Phe Gln Glu Thr Tyr
15          50          55          60
Asp Phe Ile Val Lys His Lys Phe Ser Glu Leu His Val Phe Pro Tyr
65          70          75          80
Ser Pro Arg Ile Gly Thr Pro Ala Ala Arg Met Asp Asp Gln Ile Asp
20          85          90          95
Glu Glu Ile Lys Asn Glu Arg Val His Lys Val Asn Tyr Ala
          100          105          110

```

(2) INFORMATION FOR SEQ ID NO:392:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 76 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:392:

```

Val Phe Val Val Glu Ala Leu Val Lys Thr Arg Glu Gly His Gly Asn
 1           5           10           15
Leu Glu Leu Leu Asp Lys Glu Val Ala Thr Pro Leu Asp Asp Lys Val
          20          25          30
Lys Ile Lys Val His Tyr Ala Gly Ile Cys Gly Thr Asp Leu His Thr
          35          40          45
Tyr Glu Gly His Tyr Xaa Val Asn Phe Pro Val Thr Leu Gly His Glu
          50          55          60
Phe Ser Gly Xaa Ile Val Glu Val Gly Ser Arg Arg
65          70          75

```

(2) INFORMATION FOR SEQ ID NO:393:

EP 0 841 394 A2

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 98 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:393:

```

Val Thr Ser His Val Asn Ser Ile Lys Ile His Ile Gly Arg Xaa Lys
 1           5           10          15
Gln Gly Lys Thr Ile Leu Lys Asn Ile Ser Cys His Ile Pro Lys Gly
          20          25          30
Asp Lys Trp Ile Leu Tyr Gly Leu Xaa Gly Ala Gly Lys Thr Thr Leu
          35          40          45
Leu Asn Ile Leu Asn Ala Tyr Glu Pro Ala Thr Thr Gly Gly Val Asn
          50          55          60
Leu Phe Gly Lys Met Pro Gly Lys Val Gly Tyr Ser Ala Glu Thr Xaa
65          70          75          80
Arg Gln His Ile Gly Phe Val Ser His Ser Leu Leu Glu Lys Xaa Pro
          85          90          95
Arg Gly

```

(2) INFORMATION FOR SEQ ID NO:394:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 68 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:394:

```

Val Val Ile Ser Gly Ala Phe Lys Ser Thr Gly Val Tyr Gln Asp Ile
50  1           5           10          15
Asp Asp Glu Val Arg Asn Glu Ala His His Leu Leu Lys Leu Val Gly
          20          25          30
Ile Ser Ala Lys Ala Gln Gln Tyr Ile Gly Tyr Leu Ser Thr Gly Glu
55  35          40          45

```

EP 0 841 394 A2

Lys Gln Arg Val Met Ile Ala Arg Ala Leu Met Gly Gln Pro Gln Val  
 50 55 60  
 Phe Asn Phe Arg  
 65

(2) INFORMATION FOR SEQ ID NO:395:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:395:

Val Thr Xaa Gly Val Xaa Pro Lys Pro Leu His Ala Asn Leu Met Ile  
 1 5 10 15  
 Arg Thr Leu Ser Gly His Ile Gln His Ile Ala Phe Gly Pro Ile Ala  
 20 25 30  
 Ser Leu Glu Ser Ile Lys His Leu Gly Thr Asn Gly Gly Gly Phe Leu  
 35 40 45  
 Ala Gly Asn Ser Ala Thr Pro Phe Glu Asn Pro Asn Ile Trp Ser Asn  
 50 55 60  
 Phe Ile Glu Met Gly Ser Met Met Leu Leu Pro Met Ser Met Leu Phe  
 65 70 75 80  
 Leu Phe Gly Arg Met Leu Ser Arg His Gly Lys Arg Val His Arg His  
 85 90 95  
 Ala Leu Ile Leu Phe Val Ala Met Phe Phe Ile Phe Ile Ala Ile Leu  
 100 105 110  
 Thr Leu Thr Met Trp Ser Glu Tyr Arg Gly Asn Pro Ile Leu Ala Asn  
 115 120 125  
 Leu Gly Ile Tyr Gly Pro Asn Met Glu Gly Lys Glu Val Arg Phe Gly  
 130 135 140  
 Ala Gly Leu Ser Ala Leu Phe Thr Val Ile Thr Thr Ala Phe Thr Thr  
 145 150 155 160  
 Gly Ser Val Asn Asn Met His Asp Ser Leu Thr Pro Ile Gly Gly Leu  
 165 170 175  
 Gly Pro Met Val Leu Met Met Leu Asn Val Val Phe Gly Gly Glu Gly  
 180 185 190  
 Val Gly Leu Met Asn Leu Leu Ile Xaa Xaa Leu Leu Thr Val Phe Ile  
 195 200 205

# EP 0 841 394 A2

Cys Ser Leu Met Val Gly Lys Thr Pro Glu Tyr Leu Asn Met Pro Ile  
 210 215 220  
 Gly Ala Arg Glu Met Lys Cys Ile Val Leu Val Phe Leu Ile His Pro  
 225 230 235 240  
 Ile Leu Ile Leu Val Phe Ser Ala Leu Ala Phe Met Ile Pro Gly Ala  
 245 250 255  
 Ser Glu Ser Ile Thr Asn Pro Ser Phe His Gly Ile Ser Gln Val Met  
 260 265 270  
 Leu

## (2) INFORMATION FOR SEQ ID NO:396:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 102 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Protein

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:396:

Val Cys Glu Phe Ile Lys Pro Pro Thr His Val Xaa Gly Ile Ile Leu  
 1 5 10 15  
 Gly Val Val Gly Gly Tyr Val Xaa Ala Arg Asn Xaa Xaa Leu Gln Lys  
 20 25 30  
 Gln Ser Gln Ala Arg Gln Thr Ala Xaa Asp Ile Val Xaa Gln Ala His  
 35 40 45  
 Lys Glu Ala Asp Xaa Ile Lys Lys Glu Lys Leu Leu Xaa Ala Lys Glu  
 50 55 60  
 Glu Xaa Gln Leu Xaa Arg Glu Xaa Thr Glu Ala Glu Xaa Arg Glu Arg  
 65 70 75 80  
 Arg Xaa Asp Leu Xaa Arg Gln Gly Asn Pro Thr Ser Ser Lys Arg Arg  
 85 90 95  
 Lys Leu Arg Ala His Ser  
 100

## (2) INFORMATION FOR SEQ ID NO:397:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 122 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single



EP 0 841 394 A2

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:397:

Val Arg Leu His Leu Cys Leu His Gln Arg Lys Ala Phe Leu Thr Asp  
 1 5 10 15  
 Trp Ser Tyr Ile Ala Gly Asn Ile Ala Ile Val Ala Ile Ile Pro Leu  
 20 25 30  
 Leu Ile Tyr Phe Tyr Val Pro Phe Phe Lys Lys Leu Lys Val Thr Ser  
 35 40 45  
 Ala Tyr Glu Tyr Leu Glu Ala Arg Phe Gly Pro Ser Ile Arg Val Xaa  
 50 55 60  
 Gly Ser Leu Leu Phe Val Val Tyr His Leu Gly Arg Val Ala Ile Val  
 65 70 75 80  
 Ile Tyr Leu Pro Thr Leu Ala Ile Thr Ser Val Ser Asp Met Asn Pro  
 85 90 95  
 Tyr Ile Xaa Gly Ser Leu Val Gly Leu Leu Cys Ile Xaa Xaa Thr Ser  
 100 105 110  
 Xaa Gly Xaa Phe Xaa Gly Val Arg Leu Glu  
 115 120

(2) INFORMATION FOR SEQ ID NO:398:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 121 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:398:

Val Tyr Lys Ile Thr Ile Asn Glu Val Phe Asn Met Thr Glu Thr Thr  
 1 5 10 15  
 Phe Asn Pro Ile Thr Ser Leu Thr Ile Asn Asn Glu Glu Val Lys Ala  
 20 25 30  
 Lys Ala Thr Phe Met Phe Asp Lys Thr Ala Lys Lys Phe Ala Thr Glu  
 35 40 45  
 Gln Glu Asp Asn Lys Gly Arg Lys Gln Thr Ile Ser Gly Phe Thr Asn  
 50 55 60

EP 0 841 394 A2

Val Tyr Xaa Ala Leu Leu Glu Arg Asp Thr Val Ala Ile Val Asp Phe  
 65 70 75 80  
 5 Trp Glu Cys Ala Thr Ala Tyr Leu Gly Lys Ser Ala Pro Lys Arg Glu  
 85 90 95  
 Asp Ile Glu Ala Glu Ile Met Glu Ile Ile Glu Arg Glu Asn Asp Thr  
 100 105 110  
 10 Leu Ile Phe Tyr Lys Val Arg Trp Thr  
 115 120

(2) INFORMATION FOR SEQ ID NO:399:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 39 amino acids  
 (B) TYPE: amino acid  
 20 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:399:

Val Pro Ser Asn Lys Val Ala Phe Ile Pro Ser Glu Ala Gln Phe Ile  
 1 5 10 15  
 30 Gln Leu Cys Gln Asp Asp Asn Asp Val Lys Gln Ala Ser His Asn Leu  
 20 25 30  
 Tyr Asp Gly Val Thr Phe Thr  
 35 35

(2) INFORMATION FOR SEQ ID NO:400:

40 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 108 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 45 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:400:

Val Asn Gly Val Gln Asp Asn Ser Ile Gly Ser Thr Asn Glu Ser Gln  
 1 5 10 15  
 55 Phe Leu Gly Asn Tyr Ile Val Ile Lys His Ala Glu Asn Glu Tyr Ser  
 20 25 30

EP 0 841 394 A2

Leu Ile Ala His Leu Gln Gln Tyr Ser Ile Ile Val Asn Glu Gly Gln  
 35 40 45  
 5 Asn Val Lys Tyr Gly Asp Phe Leu Gly Lys Val Gly Asn Ser Gly Asn  
 50 55 60  
 Ser Thr Glu Pro His Ile His Phe Gln Val Met Asn Asp Lys Asn Ile  
 65 70 75 80  
 10 Glu Ala Cys Thr Ser Leu Lys Ile Arg Phe Leu Asn Asn Leu Glu Leu  
 85 90 95  
 Ile Lys Gly Asp Val Val Cys Gly Leu Gln Gly Glu  
 100 105

(2) INFORMATION FOR SEQ ID NO:401:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 96 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:401:

Val His Lys Glu Asn Ile Met Leu Asn Xaa Ser Ala Thr Asp Lys Glu  
 1 5 10 15  
 Ser Val Leu Xaa Gln Met Ser Asp Val Leu Phe Gln Asn Gly Phe Val  
 20 25 30  
 35 Lys Ser Thr Phe Lys Asp Ala Val Ile Asp Arg Glu Lys Glu Xaa Xaa  
 35 40 45  
 Thr Gly Leu Pro Thr His Leu Cys Ser Val Ala Ile Pro His Thr Asp  
 50 55 60  
 40 Val Glu His Ile Asn His Arg Thr Ile Gly Val Gly Val Leu Glu Lys  
 65 70 75 80  
 Glu Val Ala Val Ile Gly Met Gly Thr Ile Gly Ser Thr Gly Arg Arg  
 85 90 95

(2) INFORMATION FOR SEQ ID NO:402:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 52 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:402:

Val Arg Ala Ala Phe Gly Lys Asn Gly Gly Asn Met Gly Val Ser Gly  
 1 5 10 15  
 Ser Val Ala Tyr Met Phe Asp His Val Ala Thr Phe Gly Ile Glu Gly  
 20 25 30  
 Lys Ser Val Asp Glu Ile Leu Glu Thr Leu Met Glu Pro Arg Cys Lys  
 35 40 45  
 Met Xaa Met Met  
 50

(2) INFORMATION FOR SEQ ID NO:403:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 79 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:403:

Val Met Phe Asn Thr Thr Val Asn Ser Asp Thr Asp Val Ile Lys Tyr  
 1 5 10 15  
 Gly Arg Leu Leu Val Asp Lys Gly Ala Gln Ser Val Ile Val Ser Leu  
 20 25 30  
 Gly Gly Asp Gly Ala Ile Tyr Ile Asp Lys Glu Ile Ser Ile Lys Ala  
 35 40 45  
 Val Asn Pro Gln Gly Lys Val Val Asn Thr Val Gly Ser Gly Asp Ser  
 50 55 60  
 Thr Val Ala Gly Met Val Ala Gly Ile Ala Ser Arg Phe Asn Asp  
 65 70 75

(2) INFORMATION FOR SEQ ID NO:404:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 123 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

EP 0 841 394 A2

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:404:

```

Val Pro Ile Ser Asn Asp Ala Ile Val Glu Leu Val Ser Arg Ile Lys
 1           5           10           15
10 Pro Val Ser Glu Met Met Glu Arg Glu Thr Asp Leu Gly Val Ala Thr
    20           25           30
Glu Phe Glu Ile Ile Thr Ala Met Met Phe Leu Tyr Phe Gly Glu Ile
    35           40           45
15 His Pro Val Asp Phe Val Ile Val Glu Ala Gly Leu Gly Ile Lys Asn
    50           55           60
Asp Ser Thr Asn Val Phe Thr Pro Val Leu Ser Ile Leu Thr Ser Ile
    65           70           75           80
20 Gly Leu Asp His Thr Asp Ile Leu Gly Gly Thr Tyr Leu Asp Ile Ala
    85           90           95
Arg Asp Lys Gly Ala Ile Ile Lys Pro Asn Val Pro Val Ile Tyr Ala
    100          105          110
25 Val Lys Asn Glu Asp Ala Leu Lys Tyr Val Arg
    115          120

```

(2) INFORMATION FOR SEQ ID NO:405:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 83 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:405:

```

Val Phe Ile Gly Thr Tyr Met Ile Leu Ser Ile Arg Lys Glu Ser Asp
45  1           5           10           15
Ala Val Ile Thr Asp Thr Asp Glu Ala Leu Lys Gln Val Leu Lys Met
    20           25           30
Val Xaa Glu Asn Lys Val Ile Ser Gln Asn Asn Lys Glu Val Thr Leu
50  35           40           45
Gln Ala Asp Thr Ile Cys Val His Gly Asp Gly Glu His Ala Leu Leu
    50           55           60
Phe Val Ser Gln Ile Arg Glu Ile Leu Met Xaa Glu Gly Ile Asp Ile
55  65           70           75           80
Gln Ser Leu

```

EP 0 841 394 A2

(2) INFORMATION FOR SEQ ID NO:406:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 87 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:406:

```

Val Phe Ala Asp Arg Arg Tyr Glu Asp Asp Gly Gln Leu Val Ser Arg
 1             5             10             15
Lys Glu Ser Asp Ala Val Ile Thr Asp Thr Asp Glu Ala Leu Lys Gln
 20             25             30
Val Leu Lys Met Val Xaa Glu Asn Lys Val Ile Ser Lys Asn Asn Lys
 35             40             45
Glu Val Thr Leu Gln Ala Asp Thr Ile Cys Val His Gly Asp Gly Glu
 50             55             60
His Ala Leu Leu Phe Val Ser Gln Ile Arg Glu Ile Leu Met Lys Glu
 65             70             75             80
Gly Ile Asp Ile Gln Ser Leu
 85

```

(2) INFORMATION FOR SEQ ID NO:407:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 90 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:407:

```

Val Tyr Asp Ile Glu Val Ser Asp Tyr Lys Gly Leu Thr Tyr Lys Leu
 1             5             10             15
Glu Ala Phe Arg Gly Lys Val Ile Leu Val Val Asn Thr Ala Thr Glu
 20             25             30

```

EP 0 841 394 A2

... Cys Ile Tyr Ser Glu Gln Leu Lys Lys Leu Glu Thr Leu Xaa Gln Lys ...  
 35 40 45  
 5 Tyr Lys Asp Arg Gly Phe Val Val Leu Ser Ser Pro Asn Asn Asn Cys  
 50 55 60  
 Asp Asn Arg Gln Pro Xaa Ser Asn Glu Glu Ile Leu Lys Ile Xaa Arg  
 65 70 75 80  
 10 Xaa Glu Ile Trp Val Leu His Phe Gln Cys  
 85 90

(2) INFORMATION FOR SEQ ID NO:408:

15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 61 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: Protein

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:408:

Val Glu Ile Glu Asn Glu Pro Tyr Phe Val Gly Lys Asp Ile Ala Glu  
 1 5 10 15  
 30 Ile Leu Gly Tyr Ala Arg Ala Asp Asn Ala Ile Arg Asn His Val Asp  
 20 25 30  
 Ser Glu Asp Lys Leu Thr His Gln Phe Ser Asp Ser Arg Ser Lys Gln  
 35 35 40 45  
 Lys Cys Asn Asp Gln Ser Thr Asn Gln Asp Tyr Thr Val  
 50 55 60

40

(2) INFORMATION FOR SEQ ID NO:409:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 139 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

45

(ii) MOLECULE TYPE: Protein

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:409:

Val Arg Xaa Leu Ile Xaa Glu Ser Xaa Ile Asp Leu Ala Ser Pro Xaa  
 1 5 10 15  
 55

EP 0 841 394 A2

Asp Asn Arg Pro Phe Ile Asp Val Phe Leu Glu Pro Thr Lys Leu Tyr  
 20 25 30  
 5 Val Xaa Pro Val Leu Ala Leu Lys Lys Glu Val Ser Ile Lys Ala Met  
 35 40 45  
 Asn His Ile Thr Gly Gly Gly Phe Tyr Glu Asn Ile Pro Arg Ala Leu  
 50 55 60  
 10 Pro Ala Gly Tyr Ala Ala Arg Ile Asp Thr Thr Ser Phe Pro Thr Pro  
 65 70 75 80  
 Lys Ile Phe Asp Trp Leu Gln Gln Gln Gly Asn Ile Asp Thr Asn Glu  
 85 90 95  
 15 Met Tyr Asn Ile Phe Asn Met Gly Ile Gly Tyr Thr Val Ile Val Asp  
 100 105 110  
 Glu Lys Asp Ala Ser Thr Arg Phe Glu Asp Phe Ser Arg Thr Lys Cys  
 115 120 125  
 20 Gly Ser Leu Ser Asn Trp Ser Tyr Cys Glu Lys  
 130 135

(2) INFORMATION FOR SEQ ID NO:410:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 83 amino acids

(B) TYPE: amino acid

30

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:410:

Val Pro Ile Phe Pro Asn Lys Ala Leu Ile Glu Ser Ala Val Ala Arg  
 1 5 10 15  
 40 Gly Glu Leu Asp Glu Ser Val Phe Asn Gln Leu Val Thr Asp Met Leu  
 20 25 30  
 Leu Glu His His Tyr Asn Ile Pro Gln His Tyr Ile Asn Leu Tyr Ile  
 35 40 45  
 Asp Asn Ile Lys Thr Leu Lys Asp Val Pro Ala Ser Tyr Met Asn His  
 50 55 60  
 Ser Asn Val Asp Val Val Ala Asp Leu Leu Leu Glu Lys Ser Lys Arg  
 50 65 70 75 80  
 Xaa Tyr Gly

55

(2) INFORMATION FOR SEQ ID NO:411:



EP 0 841 394 A2

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:411:

Val Leu Phe Ile Gly Phe Phe Val Val Ile Ile Asn Thr Ile Ala Asp  
 1 5 10 15  
 Leu Leu Thr Leu Leu Asp Pro Lys Gln Arg Leu Gln Leu Gly Asn  
 20 25 30  
 Pro Thr Lys His Asn Gln Tyr Thr Ile Asp Ile Arg Lys  
 35 40 45

(2) INFORMATION FOR SEQ ID NO:412:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 101 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:412:

Val Leu Ile Glu Val His Asp Pro His Glu Leu Glu Arg Ala Xaa Xaa  
 1 5 10 15  
 Val Asn Ala Lys Leu Ile Gly Val Asn Asn Arg Asp Leu Lys Arg Phe  
 20 25 30  
 Val Thr Asn Val Glu His Xaa Asn Thr Ile Leu Glu Asn Lys Lys Pro  
 35 40 45  
 Asn His His Tyr Ile Ser Glu Ser Gly Ile His Asp Ala Ser Asp Val  
 50 55 60  
 Arg Lys Ile Leu His Ser Gly Ile Asp Gly Leu Leu Ile Gly Glu Ala  
 65 70 75 80  
 Leu Met Arg Cys Asp Asn Leu Ser Glu Phe Leu Arg Gln Leu Lys Xaa  
 85 90 95  
 Xaa Lys Val Lys Ser  
 100

EP 0 841 394 A2

(2) INFORMATION FOR SEQ ID NO:413:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 88 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:413:

Val	Ser	Asp	Glu	Pro	Xaa	Ile	Tyr	Lys	Phe	Ile	Gln	His	Asn	Xaa	Glu
1				5				10					15		
Asn	Phe	Ile	Xaa	Thr	Ala	Ser	Xaa	Ile	Met	Asp	Gly	His	Thr	Xaa	Val
		20					25					30			
Ala	Pro	Leu	Lys	Xaa	Thr	His	Lys	Leu	Pro	Cys	Ala	Phe	Cys	Ser	Tyr
		35				40					45				
Gln	Ser	Val	Cys	His	Val	Asp	Xaa	Met	Ile	Asp	Ser	Lys	Arg	Tyr	Xaa
	50				55					60					
Xaa	Val	Asp	Glu	Thr	Ile	Asn	Pro	Ile	Glu	Ala	Ile	His	Asn	Ile	Asn
65				70				75						80	
Ile	Asn	Asp	Val	Phe	Gly	Gly	Glu								
				85											

(2) INFORMATION FOR SEQ ID NO:414:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 55 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:414:

Val	Asp	His	Leu	Xaa	Val	Thr	Leu	Arg	Ile	Asp	Arg	Glu	Asn	His	Met
1				5				10					15		
Tyr	Gly	Arg	Leu	Ala	Ser	Glu	Ser	Val	Val	Glu	Asn	Met	Phe	Thr	Pro
		20					25				30				
Val	His	Asp	Asp	Asn	Leu	Lys	Asn	Glu	Val	Ile	Glu	Ala	Lys	Pro	Trp
		35				40					45				

EP 0 841 394 A2

Xaa Glu Arg Ile Thr Asn Trp  
50 55

(2) INFORMATION FOR SEQ ID NO:415:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 75 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:415:

Val Xaa Tyr Met Asp Lys Gly Leu Thr Gly His Ile Met Arg Arg Gly  
1 5 10 15  
Ile Thr Glu Ala Asp Ala Ser Ile Asn Trp Ala Leu Gly Leu Met Asn  
20 25 30  
Glu Gly Ser Gln Ile Ile Asp Asn Thr Thr Asn Leu Cys Gly Asp Arg  
35 40 45  
Thr Thr Xaa Xaa Leu Lys Ser Gly Gly Xaa Gly Thr Gly Glu Gln Lys  
50 55 60  
Ile Asn Leu Thr Ser Lys Xaa Ala Gln Ile Trp  
65 70 75

(2) INFORMATION FOR SEQ ID NO:416:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 73 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:416:

Val Tyr Lys Gln Gly Glu Pro Asn Leu Trp Thr Gly Arg Leu Asp Ser  
1 5 10 15  
Glu Thr Asp Pro Lys Lys Phe Arg His Phe Gln Thr Val Thr Phe Glu  
20 25 30  
Asp Leu Ser Lys Leu Glu Lys Ser Ser Met Pro Ser Gly Val Gly Ile  
35 40 45

EP 0 841 394 A2

Leu Gly Tyr Ala Val Gly Gln Arg Cys Cys Phe Xaa Gln Gly Ala His  
 50 55 60  
 5 Trp Cys Lys Arg Arg Thr Arg Cys Asp  
 65 70

(2) INFORMATION FOR SEQ ID NO:417:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 130 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:417:

Val His His Ile Thr His Tyr Ile Asp Gln Leu Asp Arg Phe Ser Ser  
 1 5 10 15  
 Pro Gly Asp Val Ile Lys Xaa His Ser Gly Gly Tyr His His Lys Tyr  
 20 25 30  
 Gly Thr Phe Asn Lys Leu Gly Tyr Ile Asn Glu Asn Tyr Ile Glu Leu  
 35 40 45  
 Leu Asp Val Glu Asn Asn Glu Lys Leu Lys Lys Met Ala Xaa Thr Ile  
 50 55 60  
 Glu Gly Gly Val Ala Cys Ala Thr Gln Ile Ala Gln Glu Lys Tyr Glu  
 65 70 75 80  
 Gln Gly Phe Lys Asn Met Cys Val Xaa Thr Asn Asp Ile Glu Ala Val  
 85 90 95  
 Lys Asn Asn Leu Gln Xaa Glu Gln Val Xaa Val Val Ala Pro Thr Gln  
 100 105 110  
 Met Glu Arg Asp Thr His Lys Asp Gly Lys Val Lys Trp Gln Leu Leu  
 115 120 125  
 Ile Leu  
 130

(2) INFORMATION FOR SEQ ID NO:418:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 92 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

EP 0 841 394 A2

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:418:

```

Val Arg Gly Ser Lys Ile Gln Thr Ile Gly Val Ile Leu Pro Ser Leu
 1           5           10           15
10 Thr Asn Pro Phe Phe Ser Ala Leu Met Gln Ser Ile His Asp His Lys
    20           25           30
Pro Ser Asp Val Asp Leu Cys Phe Leu Thr Ser Thr Ala Thr Asp Xaa
    35           40           45
15 Tyr Asp Asn Ile Lys His Leu Ile Asp Arg Gly Ile Asp Gly Leu Ile
    50           55           60
Ile Ala Gln Tyr Ile Ser Ser Pro Asp Ala Leu Asn Asn Tyr Leu Lys
65           70           75           80
20 Lys His His Val Pro Tyr Val Val Leu Asp Gln Lys
    85           90

```

(2) INFORMATION FOR SEQ ID NO:419:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 98 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:419:

```

Val Leu Xaa Asp Glu Xaa Xaa Val Asp Lys Asp Ala Leu Arg Asn Asn
40  1           5           10           15
Ala Leu Val Lys Gly Gln Phe Lys Ala Asp His Gln Tyr Gln Ile Val
    20           25           30
45 Ile Gly Pro Gly Xaa Val Asp Glu Val Tyr Lys Gln Phe Ile Asp Glu
    35           40           45
Thr Gly Ala Gln Glu Ala Ser Lys Asp Glu Ala Lys Gln Ala Ala Ala
    50           55           60
50 Lys Lys Gly Asn Pro Val Gln Arg Leu Ile Lys Leu Leu Gly Glu Ile
65           70           75           80
Phe Ile Pro Ile Leu Pro Ala Ile Val Thr Thr Gly Leu Leu Met Gly
    85           90           95
55 Ile Gln

```

EP 0 841 394 A2

(2) INFORMATION FOR SEQ ID NO:420:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 163 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:420:

```

Val Ile Gln Leu Leu Met His Met Ile His Tyr Gly Thr Ser Val His
 1             5             10             15
Ile Ile Arg Thr Gln Ser Ile Leu Asn Asp Asp Lys Val Asn Gln Val
 20             20             25             30
Cys Asp Tyr Ile Glu Leu His Phe His Glu Asp Leu Ser Leu Ser Glu
 35             40             45
Leu Ser Glu Tyr Val Gly Trp Ser Glu Ser His Leu Ser Lys Lys Phe
 50             55             60
Thr Glu Ser Leu Gly Val Gly Phe Gln His Phe Leu Asn Thr Thr Arg
 65             70             75             80
Ile Glu His Ala Lys Leu Asp Leu Thr Tyr Thr Asp Glu Thr Ile Thr
 85             90             95
Asp Ile Ala Leu Gln Asn Gly Phe Ser Ser Ala Ala Ser Phe Ala Arg
100             105             110
Thr Phe Lys His Phe Thr His Gln Thr Pro Lys Gln Tyr Arg Gly Asp
115             120             125
Arg Pro Ala Ile Thr Glu Asn Gln Gln Ser Ala Gln His Asn Tyr His
130             135             140
Asp Arg Glu Leu Ile Leu Leu Leu Asn Asp Tyr Ile Glu Glu Met Asn
145             150             155             160
Gln Phe Asn

```

(2) INFORMATION FOR SEQ ID NO:421:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 105 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

# EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:421:

```

5      Val Asn Leu Gly Gly Ser Val Thr Ser Ile Gln Pro Leu Arg Ile Asn
      1          5          10          15
      Leu Thr Ser Asn Glu Asn Phe Thr Asp Lys Asp Trp Gln Ile Thr Gly
          20          25          30
10     Ile Pro Arg Thr Leu His Ile Glu Asn Ser Thr Asn Arg Thr Asn Asn
          35          40          45
      Ala Arg Glu Arg Asn Ile Glu Leu Val Gly Asn Leu Leu Pro Gly Asp
          50          55          60
15     Tyr Phe Gly Thr Ile Arg Phe Gly Arg Lys Glu Gln Leu Phe Glu Ile
      65          70          75          80
      Arg Val Xaa Pro His Xaa Pro Gln Leu Gln Arg Gln Leu Ser Asn Xaa
          85          90          95
20     Glu Val Arg Asn Tyr Lys Ser Ala Cys
          100          105

```

(2) INFORMATION FOR SEQ ID NO:422:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 56 amino acids

(B) TYPE: amino acid

30

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:422:

```

40     Val Ile Lys Arg Leu Ile Asn Glu Thr Phe Asp Ala Asn Tyr Ile Glu
      1          5          10          15
      Val Ile Glu Gly Gly Ile Glu Glu Thr Gln Thr Leu Ile His Leu Pro
          20          25          30
45     Phe Asp Tyr Val Ser Tyr Arg Ser Glu Ile Val Gly Lys Ser Phe Ile
      35          40          45
      Lys Leu Gln Arg Asn Leu Val Leu
          50          55

```

50

(2) INFORMATION FOR SEQ ID NO:423:

(i) SEQUENCE CHARACTERISTICS:

55

(A) LENGTH: 85 amino acids

(B) TYPE: amino acid

EP 0 841 394 A2

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:423:

Val Ile Thr Phe Val Leu Val Pro Val Leu Ile Pro Thr Leu Lys Arg  
 1 5 10 15  
 Met Lys Phe Gly Gln Ser Ile Arg Glu Glu Gly Pro Gln Ser His Met  
 20 25 30  
 Lys Lys Thr Gly Thr Pro Thr Met Gly Gly Leu Thr Ile Ser Ile Lys  
 35 40 45  
 Val Phe Val Ile Thr Val Phe Gly Gly Tyr Tyr Ile Cys Lys Ile Lys  
 50 55 60  
 Leu Ile Gln Ser Tyr Leu Leu Leu Phe Val Thr Asp Trp Phe Trp Val  
 65 70 75 80  
 Asn Trp Phe Tyr Arg  
 85

(2) INFORMATION FOR SEQ ID NO:424:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 49 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:424:

Val Met Xaa Asn Gly Xaa Leu Ile Glu His Gly Thr Arg Glu Ser Val  
 1 5 10 15  
 Leu His His Pro Glu His Val Tyr Thr Lys Tyr Leu Leu Ser Xaa Xaa  
 20 25 30  
 Lys Lys Xaa Asn Asp His Phe Lys His Val Met Arg Gly Asp Val His  
 35 40 45  
 Xaa

(2) INFORMATION FOR SEQ ID NO:425:

(i) SEQUENCE CHARACTERISTICS:



EP 0 841 394 A2

(A) LENGTH: 80 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:425:

Val	Pro	Thr	Ile	Asp	Ser	Val	Lys	Thr	Tyr	Gly	Leu	Val	Ser	Ile	Pro
1				5					10					15	
Gly	Met	Met	Thr	Gly	Leu	Ile	Ile	Gly	Gly	Val	Pro	Pro	Leu	Gln	Ala
			20					25					30		
Ile	Lys	Phe	Gln	Leu	Leu	Val	Val	Phe	Ile	His	Thr	Thr	Ala	Xaa	Ile
		35					40					45			
Met	Ser	Ala	Leu	Ile	Ala	Thr	Tyr	Leu	Ser	Tyr	Gly	Gln	Phe	Phe	Asn
	50					55				60					
Ala	Arg	His	Gln	Leu	Val	Ala	Arg	Asn	Thr	Asp	Val	Lys	Ser	Glu	Ser
65				70						75				80	

(2) INFORMATION FOR SEQ ID NO:426:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 100 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:426:

Val	Met	Pro	Xaa	Val	Phe	Trp	Xaa	Pro	Phe	Pro	Xaa	Ile	Phe	Ile	Gly
1				5					10					15	
Thr	Ala	Leu	Pro	Leu	Ala	Gly	Thr	Val	Ala	Thr	Gly	Ala	Ile	His	Phe
			20					25					30		
Thr	Ala	Asn	Glu	Val	Ile	Pro	Ile	Gly	Xaa	Met	Leu	Xaa	Asn	Asn	Gly
		35					40					45			
Leu	Ile	Ala	Ile	Asn	Leu	Ala	Tyr	Gln	Asn	Leu	Asp	Arg	Ala	Phe	Val
	50					55				60					
Gln	Asp	Gly	Thr	Asn	Ile	Glu	Ser	Lys	Leu	Ser	Leu	Ala	Ala	Thr	Pro
65				70						75				80	
Lys	Leu	Ala	Ser	Lys	Gly	Ala	Ile	Arg	Glu	Ser	Ile	Arg	Leu	Ala	Ile
				85						90				95	

# EP 0 841 394 A2

... Gly Ala Asn Asn  
100

5 (2) INFORMATION FOR SEQ ID NO:427:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 amino acids
- 10 (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:427:

20 Val Lys Val Gly Lys Ser Met Glu Gly Leu Asn His Arg Arg Asn Thr  
1 5 10 15  
Glu Lys Glu Glu Thr Thr Gln Thr Gln Ser Val Ala Pro Asn Thr Gly  
20 25 30  
25 Glu Glu Gly Met Ser Ser Gly Lys  
35 40

30 (2) INFORMATION FOR SEQ ID NO:428:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 62 amino acids
- 35 (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:428:

45 Val Gln Phe Leu Asn Gln Trp Leu Gly Asp Thr Leu Ala Phe Lys Val  
1 5 10 15  
Lys His Met Leu Arg Gln Arg Val Ile Tyr Lys Asn Asn Gly His Pro  
20 25 30  
50 Ile Gly Glu Gln Met Thr Ile Leu Thr Glu Asn Ile Asp Gly Leu Ala  
35 40 45  
Pro Phe Tyr Lys Ser Tyr Leu Pro Gln Val Val Gln Ile Asp  
50 55 60

55 (2) INFORMATION FOR SEQ ID NO:429:

EP 0 841 394 A2

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 116 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:429:

```

Val Val Ile Xaa Ile Ala Leu Ile Ser Xaa Val Phe Leu Pro Gly Ile
1           5           10           15
Phe Tyr Val Ile Ser Cys Val Val Gly Tyr Leu Cys Phe Gly Leu Gly
          20           25           30
Leu Gly Ile Tyr Ala Thr Pro Ser Thr Asp Thr Ala Ile Ser Asn Ala
          35           40           45
Pro Leu Asp Lys Val Gly Val Ala Ser Gly Ile Tyr Lys Met Ala Ser
          50           55           60
Ser Leu Gly Gly Ala Phe Gly Val Ala Ile Ser Gly Ala Val Tyr Val
65           70           75           80
Gly Ala Val Ala Ala Thr Ser Ile His Thr Gly Ala Met Ile Ala Leu
          85           90           95
Trp Val Asn Val Leu Met Gly Ile Met Ala Phe Ile Ala Ile Leu Phe
          100          105          110
Gly Ala Xaa Ser
          115

```

(2) INFORMATION FOR SEQ ID NO:430:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 107 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:430:

```

Val Glu Leu Gly Tyr Leu Phe Arg Asn Tyr Arg Leu Phe Asn Met Asp
1           5           10           15
Gly Leu Ala Leu Lys Leu Asn Leu Ser Ser Cys Leu Phe Ser Arg Lys
          20           25           30

```

EP 0 841 394 A2

Ser Met Ile Tyr Phe Asn Ile Gly Gln Ile Ile Ala Asn Ile Ile Cys  
 35 40 45  
 Trp Ala Leu Ile Ala Pro Thr Leu Asp Ile Leu Ile Tyr Asn Glu Pro  
 5 50 55 60  
 Ala Asn Lys Val Tyr Thr Gln Gly Val Ile Ser Ala Val Leu Asn Ile  
 65 70 75 80  
 Ile Ser Val Gly Ile Ile Gly Thr Ile Leu Leu Lys Ala Tyr Ala Ser  
 10 85 90 95  
 Ser Gln Ile Lys Lys Gly Ser Leu Arg Lys Glu  
 100 105

(2) INFORMATION FOR SEQ ID NO:431:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 137 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:431:

Val Tyr Ala Gly Glu Asn Phe Met Ile Leu Leu Ala Ser Arg Leu Asp  
 1 5 10 15  
 Ala Val Val Tyr Ser Leu Gly Leu Ala Arg Thr Arg Arg Gln Ala Arg  
 20 25 30  
 Gln Leu Val Asn His Gly His Ile Leu Val Asp Gly Lys Arg Val Asp  
 35 40 45  
 Ile Pro Ser Tyr Ser Val Lys Pro Gly Gln Thr Ile Ser Val Arg Glu  
 50 55 60  
 Lys Ser Gln Lys Leu Asn Ile Ile Val Glu Ser Val Glu Ile Asn Asn  
 65 70 75 80  
 Phe Val Pro Glu Tyr Leu Asn Phe Asp Ala Asp Ser Leu Thr Gly Thr  
 85 90 95  
 Phe Val Arg Leu Pro Glu Arg Ser Glu Leu Pro Ala Glu Ile Asn Glu  
 100 105 110  
 Gln Leu Ile Val Glu Tyr Ser Cys Gln Asp Asn Thr Phe Ile Thr Ile  
 115 120 125  
 Thr His Asn Cys Gly Cys Phe Phe Tyr  
 130 135

(2) INFORMATION FOR SEQ ID NO:432:

EP 0 841 394 A2

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 104 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:432:

```

Val His Leu Val Arg Met Val Tyr Lys Gly Lys Ile Ser Asp His Tyr
 1             5             10             15
Leu Ser Val Trp Glu Lys Ala Thr Thr Tyr Gln Met Tyr His Gly Leu
          20             25             30
Ala Leu Leu Ile Ile Gly Val Ile Ser Gly Thr Thr Ser Ile Asn Val
          35             40             45
Asn Trp Ala Gly Trp Leu Ile Phe Ala Gly Ile Ile Phe Phe Ser Gly
          50             55             60
Ser Leu Tyr Ile Leu Val Leu Thr Gln Ile Lys Val Leu Gly Ala Ile
65             70             75             80
Thr Pro Ile Gly Gly Val Leu Phe Ile Ile Gly Trp Ile Met Leu Ile
          85             90             95
Ile Ala Thr Phe Lys Phe Ala Gly
          100

```

(2) INFORMATION FOR SEQ ID NO:433:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 65 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:433:

```

Val Leu Lys Leu Phe Gln Met Arg Ser Pro Ile Phe Arg Glu Pro Ser
 1             5             10             15
Ala Asn Asn Ala Val Lys Thr Leu Ile Ala Met Gly Ser Ile Leu Ala
          20             25             30
Phe Leu Leu Val Gly Ile Gly Gly Leu Ala Tyr Val Tyr Gly Ile Met
          35             40             45

```

EP 0 841 394 A2

... Pro Gln Thr Glu Thr Thr Val Leu Ser Gln Leu Ala Met Gln Ile Phe ...  
 50 55 60

Gly  
 65

(2) INFORMATION FOR SEQ ID NO:434:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 68 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:434:

Val Met Arg Gly Thr Ile Ile Ile Pro Thr Thr Lys Pro Gly Leu Ile  
 1 5 10 15  
 Ala Leu Asn Ser Pro Arg Pro Asp Met Lys Asp Leu Asn Thr Gly Val  
 20 25 30  
 Thr Lys Val Asn Ala Lys Lys Pro Asn Thr Ile Val Gly Ile Pro Ala  
 35 40 45  
 Lys Ile Ser Asn Ile Gly Leu Ile Met Arg Arg Ala Arg Ala Leu Ala  
 50 55 60  
 Tyr Ser Leu Lys  
 65

(2) INFORMATION FOR SEQ ID NO:435:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:435:

Val Ser Asp Thr Ala Val Met Ile Ala Trp Leu Val Lys Ile Pro His  
 1 5 10 15  
 Ser Thr Met Pro Ile Leu Gly Thr Ser Gln Leu Lys Arg Xaa Asp Gln  
 20 25 30

EP 0 841 394 A2

Ala Ile Glu Gly Leu Gln Leu Asn Leu Asp Asp Gln Val Val Gly  
35 40 45

(2) INFORMATION FOR SEQ ID NO:436:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:436:

Val Asp Ile Pro Leu Leu Phe Glu Asn Glu Leu Glu Asn Thr Val Asp  
1 5 10 15  
Glu Val Trp Val Val Tyr Thr Ser Glu Ser Ile Gln Met Asp Arg Leu  
20 25 30  
Met Xaa Arg Xaa Asp Leu Ser Leu Glu Asp Ala Lys Ala Arg Val Tyr  
35 40 45  
Xaa Pro Asn Phe Tyr  
50

(2) INFORMATION FOR SEQ ID NO:437:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 111 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:437:

Val Met Gln Val Phe Thr Trp Gln Gln Tyr Val Glu Ile Val Val Asn  
1 5 10 15  
Glu Gly Arg Asp Ala Ala Asn Ala Ala Gln Glu Lys Ala Val Lys Glu  
20 25 30  
Gly Lys Ile Ile Ile Lys Asp Ser Ile Ala Asp Ile Phe Leu Gln Gln  
35 40 45  
Ile Leu Thr Arg Pro Ala Glu His Asp Val Val Ala Thr Met Asn Leu  
50 55 60

EP 0 841 394 A2

Asn Gly Asp Tyr Ile Ser Asp Ala Leu Ala Ala Gln Val Gly Xaa Ile  
 65 70 75 80  
 Gly Ile Ala Pro Gly Ala Asn Ile Asn Tyr Glu Thr Gly His Ala Ile  
 5 85 90 95  
 Phe Glu Ala Thr His Gly Leu Xaa Ser Lys Ile Cys Arg Phe Lys  
 100 105 110

(2) INFORMATION FOR SEQ ID NO:438:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:438:

Val Ser Met Phe Ile Thr Gly Met Gln Tyr Gly Asp Lys Val Ala Val  
 1 5 10 15  
 His Val Ser Arg Gly Ala Val Phe Gly Met Thr Gly Val Leu Val Val  
 20 25 30  
 Phe

(2) INFORMATION FOR SEQ ID NO:439:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 105 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:439:

Val Val Thr Asp Gln Leu Leu Ala Phe Phe Asn Asn Arg Tyr Trp Arg  
 1 5 10 15  
 Ser Gln Phe Asn Pro Arg Gly Gly Trp Ser Pro Ser Gly Pro Arg Arg  
 20 25 30  
 Tyr Ala Asn Gly Gly Leu Ile Thr Lys His Gln Leu Xaa Glu Val Gly  
 35 40 45



EP 0 841 394 A2

... Glu Gly Asp Lys Gln Glu Met Val Ile Pro Leu Thr Arg Arg Lys Arg ...  
 50 55 60  
 5 Ala Ile Gln Leu Thr Glu Gln Val Met Arg Ile Ile Gly Met Asp Gly  
 65 70 75 80  
 Xaa Pro Asn Asn Ile Thr Val Asn Asn Asp Thr Ser Thr Val Glu Lys  
 85 90 95  
 10 Leu Phe Glu Thr Asn Cys Tyr Val Lys  
 100 105

(2) INFORMATION FOR SEQ ID NO:440:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 109 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:440:

Val Ala Gly Ala Leu Asp Val Asn Gln Val Asn Val Ser Glu Asn Asp  
 1 5 10 15  
 30 Asn Ala Asn Gln Pro His Ser Val Leu Leu Ile Asp Thr Gln Ala Val  
 20 25 30  
 Asp Glu Asn Asn Ser Glu Leu Asn Gln Val Gly Thr Ser Thr Lys Ala  
 35 35 40 45  
 Gln Ile Ala Phe Cys Ile Asp Val Arg Ser Glu Pro Phe Arg Arg His  
 50 55 60  
 Ile Glu Ala Ala Gly Pro Phe Glu Thr Ile Gly Ile Ala Gly Phe Phe  
 40 65 70 75 80  
 Gly Leu Pro Ile Gln Lys Asp Ala Val Asp Glu Gln Phe Lys His Asp  
 85 90 95  
 45 Ser Leu Pro Val Met Ser Arg Ala Ala Gly Ile Ser His  
 100 105

(2) INFORMATION FOR SEQ ID NO:441:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

EP 0 841 394 A2

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:441:

Val Thr Phe Gly Val Thr Gly Thr Asn Gly Lys Thr Ser Ile Ala Thr  
 1 5 10 15  
 Asp Asp Ser Phe Asn Ser Xaa Lys Val Thr Lys Lys  
 20 25

(2) INFORMATION FOR SEQ ID NO:442:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 110 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:442:

Val Asp Gly Thr Ala Asn Gly Val Gly Ser Thr Leu Xaa Leu Asn Glu  
 1 5 10 15  
 Ser Leu Asp Gln Phe Ile Leu Leu Ile Phe Tyr Gly Thr Phe Pro Gly  
 20 25 30  
 Gly Asp Phe Thr Glu Phe Gly Ser Pro Phe Gly Gly Gly Lys Ile Ser  
 35 40 45  
 Leu Asn Pro Ser Asn Leu Pro Asp Gly Asp Gly Asn Gly Gly Gly Val  
 50 55 60  
 Tyr Glu Phe Gly Leu Thr Lys Ser Ser Arg Thr Ser Leu Thr Ile Ser  
 65 70 75 80  
 Asn Asp Val Tyr Phe Asp Leu Gly Ser Gln Arg Gly Ser Gly Ala Asn  
 85 90 95  
 Ala Asn Arg Gly Thr Ile Asn Lys Ile Ile Gly Val Arg Lys  
 100 105 110

(2) INFORMATION FOR SEQ ID NO:443:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 108 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

EP 0 841 394 A2

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:443:

```

Val Leu Asp Asn Asp Leu Asn Gln Val Thr Leu Ala Asp Tyr Ala Gly
 1           5           10          15
Lys Lys Lys Leu Ile Ser Val Val Pro Ser Ile Asp Thr Gly Val Cys
10          20          25          30
Asp Gln Gln Thr Arg Lys Phe Asn Ser Glu Ala Ser Lys Glu Glu Gly
          35          40          45
Ile Val Leu Thr Ile Ser Ala Asp Leu Pro Phe Ala Gln Lys Arg Trp
15          50          55          60
Cys Ala Ser Ala Gly Leu Asp Asn Val Ile Thr Leu Ser Asp His Arg
65          70          75          80
Asp Leu Ser Phe Gly Glu Asn Tyr Gly Val Val Met Glu Arg Thr Ser
20          85          90          95
Cys Arg Ile Arg His Glu Leu Val Gln Tyr Leu Tyr
          100          105

```

(2) INFORMATION FOR SEQ ID NO:444:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 133 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:444:

```

Val Ser Gln Glu Arg Tyr Ser Arg Gln Ile Leu Phe Lys Gln Ile Gly
 1           5           10          15
Glu Ile Gly Gln Ser Lys Ile Asn Gln Lys Cys Ala Leu Ile Ile Gly
45          20          25          30
Met Gly Ala Leu Gly Thr His Val Ala Glu Gly Leu Val Arg Ala Gly
          35          40          45
Ile Ala Lys Leu Ile Ile Val Asp Arg Asp Tyr Ile Glu Phe Ser Asn
50          50          55          60
Leu Gln Arg Gln Thr Leu Phe Thr Glu Glu Asp Ala Leu Lys Met Met
65          70          75          80
Pro Lys Val Val Ala Ala Lys Lys His Leu Leu Ala Leu Arg Ser Asp
55          85          90          95

```

EP 0 841 394 A2

Val Asp Ile Asp Gly Cys Ile Ala His Val Asp Tyr Tyr Phe Leu Gly  
 100 105 110  
 Asn Thr Trp Thr Gly Arg Trp Thr Leu Leu Leu Met Gln Pro Ile Thr  
 115 120 125  
 Leu Lys His Asp Asn  
 130

(2) INFORMATION FOR SEQ ID NO:445:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 144 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:445:

Val Thr Leu Met Lys Leu Val Phe Val Ala Arg Ala Gly Asn Met Ala  
 1 5 10 15  
 Gln Ala Ile Phe Thr Gly Ile Ile Asn Ser Ser Asn Leu Asp Ala Asn  
 20 25 30  
 Asp Ile Tyr Leu Thr Asn Lys Ser Asn Glu Gln Ala Leu Lys Ala Phe  
 35 40 45  
 Ala Glu Lys Leu Gly Val Asn Tyr Ser Tyr Asp Asp Ala Thr Leu Leu  
 50 55 60  
 Lys Asp Ala Asp Tyr Val Phe Leu Gly Thr Lys Pro His Asp Phe Asp  
 65 70 75 80  
 Ala Leu Ala Thr Arg Ile Lys Pro His Ile Thr Lys Asp Xaa Cys Phe  
 85 90 95  
 Met Ser Ile Met Ala Gly Ile Pro Thr Asp Tyr Ile Xaa Gln Gln Leu  
 100 105 110  
 Glu Cys Gln Asn Pro Xaa Ala Arg Ile Met Pro Xaa Thr Xaa Ala Xaa  
 115 120 125  
 Val Gly His Ser Val Thr Gly Ile Ser Phe Ser Asn Asn Phe Glu Pro  
 130 135 140

(2) INFORMATION FOR SEQ ID NO:446:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single

EP 0 841 394 A2

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:446:

Val His Arg Ser Leu Met Lys Ile Leu Ile Asn Val Leu Glu Gln Phe  
1 5 10 15  
Asp Asn Ala Val Val Leu Phe Met Gly Ala Gly Asp Ile Gln Lys Leu  
20 25 30  
Gln Asn Ala Tyr Leu Asp Lys Leu Gly Met Lys Asn Ala Phe  
35 40 45

(2) INFORMATION FOR SEQ ID NO:447:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 64 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:447:

Val Phe Glu Asn Glu Pro Leu Lys Pro Asn His Glu Leu Tyr Glu Leu  
1 5 10 15  
Glu Asn Val Thr Ile Thr Ala His Ile Thr Gly Asn Asp Tyr Glu Ala  
20 25 30  
Lys Tyr Asp Leu Leu Asp Ile Phe Lys Asn Asn Leu Val Asn Phe Leu  
35 40 45  
Asn Lys Asn Gly Leu Ile Glu Asn Glu Val Asp Ala Lys Lys Gly Tyr  
50 55 60

(2) INFORMATION FOR SEQ ID NO:448:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 86 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

# EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID.NO:448:

```

5      Val Lys Glu Ile Thr Lys Leu Asn Gly Leu Cys Leu Lys Trp Val Ala
      1          5          10          15
      Pro Gly Thr Arg Gly Val Pro Asp Arg Ile Ile Ile Met Pro Glu Gly
          20          25          30
10     Lys Thr Tyr Phe Val Glu Met Lys Gln Glu Lys Gly Lys Leu His Pro
          35          40          45
      Leu Gln Lys Tyr Val His Arg Gln Phe Glu Asn Arg Asp His Lys Val
          50          55          60
15     Tyr Val Leu Trp Asn Lys Glu Gln Val Asn Thr Phe Ile Arg Met Val
      65          70          75          80
      Gly Gly Thr Phe Gly Asp
          85

```

(2) INFORMATION FOR SEQ ID NO:449:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 115 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:449:

```

35     Val Trp Gly Ile Val Ile Leu Gly Gly Tyr Glu Gln Phe Ile Lys Ser
      1          5          10          15
      His Leu Arg Lys Ile Tyr Ile Asp Gly Val Ser Asn Met Gln Glu His
          20          25          30
40     Val Val Val Thr Leu Asp Gly Lys Asp Tyr Leu Val Glu Pro Gly Thr
          35          40          45
      Asn Leu Leu Glu Phe Ile Lys Ser Gln Asp Thr Phe Val Pro Ser Ile
          50          55          60
45     Cys Tyr Asn Glu Ser Met Gly Pro Ile Gln Thr Cys Asp Thr Cys Thr
      65          70          75          80
      Val Glu Ile Asp Gly Lys Ile Glu Arg Ser Cys Ser Thr Val Ile Asp
          85          90          95
50     Arg Pro Met Thr Val Asn Thr Val Asn Asn Asp Val Lys Asp Ala Gln
          100          105          110
      Lys Glu Pro
          115
55

```

EP 0 841 394 A2

(2) INFORMATION FOR SEQ ID NO:450:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:450:

```

Val Leu Thr Met Ile Ser Tyr Leu Leu Phe Leu Leu Ser Gly Leu Ala
 1             5             10             15
Asn Gly Leu Ile Asn Met Asn Lys Glu Gly Ile Asp Lys Trp Gln Ala
20             20             25             30
Asp Ala Ile Xaa Leu Asn Lys Asp Ala Asn Gln Thr Val Gln Xaa Ser
35             40             45
Cys Phe
25             50

```

(2) INFORMATION FOR SEQ ID NO:451:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 174 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:451:

```

Val Lys Lys Val Val Lys Tyr Leu Ile Ser Leu Ile Leu Ala Ile Ile
45 1             5             10             15
Ile Val Leu Phe Val Gln Thr Phe Val Ile Val Gly His Val Ile Pro
20             25             30
Asn Asn Asp Met Ser Pro Thr Leu Asn Lys Gly Asp Arg Val Ile Val
50 35             40             45
Asn Lys Ile Lys Val Thr Phe Asn Gln Leu Asn Asn Gly Asp Ile Ile
50             55             60
Thr Tyr Arg Arg Gly Asn Glu Ile Tyr Thr Ser Arg Ile Ile Ala Lys
55 65             70             75             80

```

EP 0 841 394 A2

.. Pro Gly Gln Ser Met Ala Phe Arg Gln Gly Gln Leu Tyr Arg Asp Asp ..  
 85 90 95  
 5 Arg Pro Val Asp Ala Ser Tyr Ala Lys Asn Arg Lys Ile Lys Asp Phe  
 100 105 110  
 Ser Leu Arg Asn Phe Lys Glu Leu Asp Gly Asp Ile Ile Pro Pro Asn  
 115 120 125  
 10 Asn Phe Val Val Leu Asn Asp His Asp Asn Asn Gln His Asp Ser Arg  
 130 135 140  
 Gln Phe Gly Leu Ile Asp Lys Lys Asp Ile Ile Gly Asn Ile Ser Leu  
 145 150 155 160  
 15 Arg Tyr Tyr Pro Phe Ser Lys Trp Thr Ile Gln Phe Lys Ser  
 165 170

(2) INFORMATION FOR SEQ ID NO:452:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 146 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:452:

Val Ser Lys Leu Lys Lys Glu Leu Leu Glu Trp Ile Ile Ser Ile Ala  
 1 5 10 15  
 Val Ala Phe Val Ile Leu Phe Ile Val Gly Lys Phe Ile Val Thr Pro  
 20 25 30  
 Tyr Thr Ile Lys Gly Glu Ser Met Asp Pro Thr Leu Lys Asp Gly Glu  
 35 40 45  
 Arg Val Ala Val Asn Ile Ile Gly Tyr Lys Thr Gly Gly Leu Glu Lys  
 50 55 60  
 Gly Asn Val Val Val Phe His Ala Asn Lys Asn Asp Asp Tyr Val Lys  
 65 70 75 80  
 Arg Val Ile Gly Val Pro Gly Asp Lys Val Glu Tyr Lys Asn Asp Thr  
 85 90 95  
 Leu Tyr Val Asn Gly Lys Lys Gln Asp Glu Pro Tyr Leu Asn Tyr Asn  
 100 105 110  
 Leu Lys His Lys Gln Gly Asp Tyr Ile Thr Gly Thr Phe Gln Val Lys  
 115 120 125  
 Asp Leu Pro Glu Cys Glu Ser Gln Ile Lys Cys Gln Ser Lys Arg Val  
 130 135 140



# EP 0 841 394 A2

... Asn Ile .  
145

5

## (2) INFORMATION FOR SEQ ID NO:453:

### (i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 107 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

15

### (ii) MOLECULE TYPE: Protein

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:453:

20

Val	Glu	Arg	Asn	Gln	Lys	Val	Pro	Xaa	Gly	Pro	Glu	Asn	Pro	Glu	Lys
1				5					10					15	
Pro	Ser	Arg	Pro	Thr	His	Pro	Xaa	Gly	Pro	Val	Asn	Pro	Asn	Asn	Pro
			20					25					30		
Gly	Leu	Ser	Xaa	Asp	Arg	Ala	Lys	Pro	Asn	Gly	Pro	Gly	Pro	Phe	Asn
		35				40					45				
Trp	Ile	Lys	Met	Ile	Lys	Val	Lys	Lys	Ser	Lys	Ile	Ala	Lys	Glu	Ser
	50				55				60						
Val	Ala	Asn	Gln	Glu	Lys	Lys	Arg	Ala	Glu	Leu	Pro	Lys	Thr	Gly	Leu
	65				70				75					80	
Glu	Ser	Thr	Gln	Lys	Gly	Leu	Ile	Phe	Ser	Ser	Ile	Ile	Gly	Ile	Ala
			85					90					95		
Gly	Leu	Met	Leu	Leu	Ala	Arg	Arg	Arg	Lys	Asn					
			100					105							

40

## (2) INFORMATION FOR SEQ ID NO:454:

### (i) SEQUENCE CHARACTERISTICS:

45

- (A) LENGTH: 54 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

50

### (ii) MOLECULE TYPE: Protein

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:454:

55

Val	Thr	His	Val	Asp	Val	Asp	Glu	Arg	Leu	Ile	Asp	Phe	Gln	Ile	Val
1			5					10					15		

EP 0 841 394 A2

Gly Met Pro Leu Pro Lys Asn Asp Arg Ser Gln Arg Pro Ala Arg Gly  
 20 25 30  
 Lys Thr Ile Gln Ala Lys Thr Arg Gly Lys Ser Leu Asp Lys Ser Lys  
 35 40 45  
 Ser Asp Asp Lys Gly Arg  
 50

(2) INFORMATION FOR SEQ ID NO:455:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 62 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:455:

Val Phe His Ile Leu Gln Asn Ile Gly Met Thr Ile Gln Leu Leu Pro  
 1 5 10 15  
 Ile Thr Gly Ile Pro Leu Pro Phe Ile Ser Tyr Gly Gly Ser Ala Leu  
 20 25 30  
 Trp Ser Met Met Thr Gly Ile Gly Ile Val Leu Ser Ile Tyr Tyr His  
 35 40 45  
 Glu Pro Lys Arg Tyr Val Asp Leu Tyr His Pro Lys Ser Asn  
 50 55 60

(2) INFORMATION FOR SEQ ID NO:456:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 74 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:456:

Val Thr Xaa Arg Glu Val Val Asn His Ile Glu Gln Thr Ile Xaa Gln  
 1 5 10 15  
 Tyr Gly Val Asn Glu Met Ser Phe Asp Thr Met Val Leu Phe Gly Asp  
 20 25 30

EP 0 841 394 A2

His Xaa Ala Ser Pro His Gly Thr Pro Gly Asp Arg Arg Leu Lys Ser  
 35 40 45  
 Asn Glu Tyr Val Leu Phe Asp Leu Gly Val Ile Tyr Glu His Tyr Cys  
 50 55 60  
 Ser Asp Met Thr Arg Thr Ile Lys Phe Trp  
 65 70

(2) INFORMATION FOR SEQ ID NO:457:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:457:

Val Gln Val Gly Asp Gly Pro Xaa Gln Arg Glu Ile Val Phe Pro Asn  
 1 5 10 15  
 Ser Thr Asp Xaa Xaa Ser Trp Gly Gly Xaa Thr Lys  
 20 25

(2) INFORMATION FOR SEQ ID NO:458:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 70 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:458:

Val Gly Pro Asn Thr Glu Ala Glu Phe Arg Lys Glu Ile Gln Leu Pro  
 1 5 10 15  
 Gly Asn Ala Ser Trp Gly Val Gly Pro Asn Thr Glu Xaa Ile Gly Phe  
 20 25 30  
 Pro Ile Ser Ser Asp Asn Ala Ser Trp Gly Val Gly Pro Asn Lys Glu  
 35 40 45  
 Asn Phe Glu Lys Lys Phe Tyr Arg Gln Cys Glu Leu Gly Cys Gly Pro  
 50 55 60

EP 0 841 394 A2

... Gln His Arg Ser Leu Arg  
65 70

(2) INFORMATION FOR SEQ ID NO:459:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:459:

Val Thr Arg Leu Phe Tyr Met Phe Asp Lys Glu Thr Met Ile Ala Ser  
1 5 10 15  
Met Gly Ile Gly Gly Leu Gly Asn Ala Ala Leu Phe Thr Arg Phe  
20 25 30

(2) INFORMATION FOR SEQ ID NO:460:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:460:

Val Ser Ala Ile Val Xaa Tyr Xaa Ser Phe Asp Cys Xaa Ile Arg Lys  
1 5 10 15  
Phe Lys Leu Lys Thr Pro Gly Arg Glu Asp Glu Glu Thr Glu Ile Arg  
20 25 30  
Asn Ser Ser Val Ala Lys Leu Pro Xaa Asp Val Leu Asp Xaa Met Gly  
35 40 45  
Trp Lys Arg Lys His  
50

(2) INFORMATION FOR SEQ ID NO:461:

(i) SEQUENCE CHARACTERISTICS:

EP 0 841 394 A2

(A) LENGTH: 136 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:461:

```

Val Leu Arg Met Thr Glu Gln Pro Ser Tyr Tyr Ser Ile Ile Thr Ala
1           5           10           15
Asn Val Arg Tyr Asp Asn Arg Leu Thr Asp Ser Glu Lys Leu Leu Phe
           20           25           30
Ala Glu Ile Thr Ser Leu Ser Asn Lys Tyr Gly Tyr Cys Thr Ala Ser
20          35          40          45
Asn Gly Tyr Phe Ala Lys Leu Tyr Glu Val Thr Lys Val Thr Val Ser
           50           55           60
Arg Arg Ile Ala Asn Leu Lys Glu Cys Gly Tyr Leu His Val Glu Ile
25          65          70          75          80
Ile Arg Asn Gly Asn Glu Ile Lys Gln Arg Lys Leu Tyr Pro Leu Thr
           85           90           95
Glu Met Ile Arg Pro Ile Asn Thr Asn Asp Asn Thr Pro Ile Asn Asn
30          100         105         110
Ser Val Asn Thr Pro Ile Ile Thr Asn Val Lys Glu Thr Ile Gln Val
           115         120         125
Leu Ile Ile Gln Val Ile Thr Ile
35          130         135

```

(2) INFORMATION FOR SEQ ID NO:462:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 290 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:462:

```

Val Gly Phe Xaa Ala Ala Ile Arg Met Asp Lys Ser Gly Tyr Ser Xaa
1           5           10           15
Ser Leu Tyr Glu Gln Asn Thr His Ile Gly Gly Lys Val Asn Xaa His
20          25          30

```

EP 0 841 394 A2

```

5  Glu Ser Asp Gly Phe Gly Phe Asp Leu Gly Pro Ser Ile Leu Thr Met
    35                                40                                45
    Pro Tyr Ile Cys Glu Lys Leu Phe Glu Tyr Ser Lys Lys Gln Met Ser
    50                                55                                60
    Asp Tyr Val Thr Ile Lys Arg Leu Xaa His Gln Trp Arg Ser Phe Phe
    65                                70                                75                                80
10  Pro Asp Gly Thr Thr Ile Asp Leu Tyr Glu Gly Ile Lys Glu Thr Gly
    85                                90                                95
    Gln His Asn Ala Ile Leu Ser Lys Gln Asp Ile Glu Glu Leu Gln Asn
    100                                105                                110
15  Tyr Leu Asn Tyr Thr Arg Arg Ile Asp Arg Ile Thr Glu Lys Gly Tyr
    115                                120                                125
    Phe Asn Tyr Gly Leu Asp Thr Leu Ser Gln Ile Ile Lys Phe His Gly
    130                                135                                140
20  Pro Leu Asn Ala Leu Ile Asn Tyr Asp Tyr Val His Thr Met Gln Gln
    145                                150                                155                                160
    Ala Ile Asp Lys Arg Ile Ser Asn Pro Tyr Leu Arg Gln Met Leu Gly
    165                                170                                175
25  Tyr Phe Ile Lys Tyr Val Gly Ser Ser Ser Tyr Asp Ala Xaa Ala Val
    180                                185                                190
    Leu Ser Met Leu Phe His Met Gln Gln Glu Gln Gly Xaa Xaa Tyr Val
    195                                200                                205
30  Glu Gly Gly Ile His His Xaa Ala Asn Ala Leu Glu Lys Leu Ala Arg
    210                                215                                220
    Glu Glu Gly Val Thr Ile His Thr Gly Ala Arg Val Asp Asn Ile Lys
    225                                230                                235                                240
35  Thr Tyr Gln Arg Arg Val Thr Gly Val Arg Leu Asp Thr Gly Glu Phe
    245                                250                                255
    Val Lys Ala Asp Tyr Ile Ile Ser Asn Met Glu Val Ile Pro Thr Tyr
    260                                265                                270
40  Lys Tyr Leu Ile His Leu Gly Tyr Ser Thr Ile Lys Gln Ile Arg Glu
    275                                280                                285
    Gly Ile
    290

```

(2) INFORMATION FOR SEQ ID NO:463:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 224 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:463:

5

10

15

20

25

30

35

40

```

Val Arg Thr Ser Arg Gln Ser Pro Phe Ala Gln Arg Ile Glu Gln Gln
 1           5           10           15
Gln Lys Arg Leu Asn Leu Pro Asp Leu Pro Thr Thr Thr Ile Gly Ser
20           25           30
Phe Pro Gln Ser Arg Glu Val Arg Lys Tyr Arg Ala Asp Trp Lys Asn
35           40           45
Lys Arg Ile Thr Asp Glu Ala Tyr Glu Thr Phe Leu Lys Asn Glu Ile
50           55           60
Ala Arg Trp Ile Lys Ile Gln Glu Asp Ile Gly Leu Asp Val Leu Val
65           70           75           80
His Gly Glu Phe Glu Arg Asn Asp Met Val Glu Phe Phe Gly Glu Lys
85           90           95
Leu Gln Gly Phe Leu Val Thr Lys Phe Gly Trp Val Gln Ser Tyr Gly
100          105          110
Ser Arg Ala Val Lys Pro Pro Ile Ile Tyr Gly Asp Val Lys Trp Thr
115          120          125
Ala Pro Leu Thr Val Asp Glu Thr Val Tyr Ala Gln Ser Leu Thr Asp
130          135          140
Lys Pro Val Lys Gly Met Leu Thr Gly Pro Val Thr Ile Leu Asn Trp
145          150          155          160
Ser Phe Glu Arg Val Asp Leu Pro Arg Lys Val Ala Gln Asp Gln Ile
165          170          175
Ala Leu Ala Ile Asn Glu Glu Val Leu Ala Leu Glu Ala Ala Gly Ile
180          185          190
Lys Val Ile Gln Val Asp Glu Pro Ala Leu Arg Glu Gly Leu Pro Leu
195          200          205
Arg Ser Glu Tyr His Glu Gln Tyr Leu Lys Asp Ala Gly Phe Ile Ile
210          215          220

```

(2) INFORMATION FOR SEQ ID NO:464:

45

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 118 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:464:

EP 0 841 394 A2

Val Lys Met Met Pro Arg Lys Phe Arg Val Leu Gln Ile Gly Gly Asp  
 1 5 10 15  
 5 Asp Leu Glu Pro Ile Phe Gln His Lys Lys Gly Val Ser Trp Asp Tyr  
 20 25 30  
 Phe Asp Ile Gly Leu Phe Glu Phe Asp Ser Gly Tyr Val Glu Ala Ile  
 35 40 45  
 10 Glu Ala Ile Val Glu Ala Glu Gly Arg Phe Asp Phe Ile Tyr Ile Gln  
 50 55 60  
 Ala Pro Tyr Ser Glu Thr Leu Thr Asn Leu Leu Gln Met Ile Ser Glu  
 65 70 75 80  
 15 Pro Tyr Asn Thr Tyr Val Asp Glu Ser Phe Trp Ser Val Glu Tyr Glu  
 85 90 95  
 Gln Asp Glu Asn Cys Pro Lys Ile Arg Cys Ser Thr Asn Tyr Ile Thr  
 100 105 110  
 20 Gly Ile Leu Glu Gly Thr  
 115

(2) INFORMATION FOR SEQ ID NO:465:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 194 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:465:

Val Arg Lys Gly Tyr Tyr Pro Ile Lys Arg Ala Ile Asp Leu Val Leu  
 1 5 10 15  
 40 Ser Ile Val Leu Leu Phe Leu Thr Phe Pro Ile Met Phe Ile Phe Ala  
 20 25 30  
 Ile Ala Ile Val Ile Asp Ser Pro Gly Asn Pro Ile Tyr Ser Gln Val  
 35 40 45  
 Arg Val Gly Lys Met Gly Lys Leu Ile Lys Ile Tyr Lys Leu Arg Ser  
 50 55 60  
 Met Cys Lys Asn Ala Glu Lys Asn Gly Ala Gln Trp Ala Asp Lys Asp  
 50 65 70 75 80  
 Asp Asp Arg Ile Thr Asn Val Gly Lys Phe Ile Arg Lys Thr Arg Ile  
 85 90 95  
 Asp Glu Leu Pro His Leu Ile Asn Val Val Lys Gly Glu Met Ser Phe  
 100 105 110



EP 0 841 394 A2

... Ile Gly Pro Arg Pro Glu Arg Pro Glu Phe Val Glu Leu Phe Ser Ser ...  
 115 120 125  
 5 Glu Val Ile Gly Phe Glu Gln Arg Cys Leu Val Thr Pro Gly Leu Thr  
 130 135 140  
 Gly Leu Ala Gln Ile Gln Gly Gly Tyr Asp Leu Thr Pro Gln His Lys  
 145 150 155 160  
 10 Leu Lys Tyr Asp Met Lys Tyr Ile His Lys Gly Ser Leu Met Met Glu  
 165 170 175  
 Leu Tyr Ile Ser Ile Arg Thr Leu Met Val Val Ile Thr Gly Glu Gly  
 180 185 190  
 15 Ser Arg

(2) INFORMATION FOR SEQ ID NO:466:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 93 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:466:

Val Lys Ala Leu Lys Leu Tyr Gly Val Glu Asp Leu Arg Tyr Glu Asp  
 1 5 10 15  
 Asn Glu Lys Pro Val Ile Glu Ser Ala Asn Asp Val Ile Val Lys Val  
 20 25 30  
 Arg Ala Thr Gly Ile Cys Gly Ser Asp Thr Ser Arg Asn Xaa Xaa Met  
 35 40 45  
 Gly Pro Tyr Ile Lys Gly Met Pro Phe Gly His Glu Phe Ser Gly Val  
 50 55 60  
 Val Asp Ala Ile Gly Ser Asp Val Thr His Val Asn Val Gly Asp Lys  
 65 70 75 80  
 Val Thr Gly Cys Pro Ala Ile Pro Cys Tyr Gln Cys Glu  
 85 90

(2) INFORMATION FOR SEQ ID NO:467:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 105 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single

EP 0 841 394 A2

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:467:

```

Val Gly Glu Ala Glu Ser Ser Leu Ser Phe Ile Gly Glu Leu Asp Asn
 1           5           10          15
Ile Asp Asp Lys Thr Tyr Gln Asp Ala Leu Val Ile Val Cys Asp Thr
          20          25          30
Ala Asn Ala Pro Arg Ile Asp Asp Glu Arg Tyr Ser Thr Gly Ser Lys
          35          40          45
Leu Ile Lys Ile Asp His His Pro Ala Val Asp Gln Tyr Gly Asp Ile
          50          55          60
Asn Leu Val Asn Thr Asn Ala Ser Ser Thr Ser Glu Ile Ile Tyr Asp
65          70          75          80
Leu Ile Ser His Phe Asn Asp Glu Ala Ile Val Asn Lys Gly Thr Ser
          85          90          95
Glu Cys Phe Ile Pro Trp Tyr Arg Arg
          100          105

```

(2) INFORMATION FOR SEQ ID NO:468:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 250 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:468:

```

Val Met Ser Leu Val Ile Val Leu Ile Lys Asp Gly Cys Val Ser Lys
 1           5           10          15
Phe Ser Leu Ile Arg Gln Gly Asn Met Ile Lys Arg Asp Xaa Pro Met
          20          25          30
Ile Pro Leu His Gln Thr Glu Glu Glu Glu Phe Tyr Thr Phe Ile Gly
          35          40          45
Gln Phe Tyr Ser Leu Asn Gln His Ile Leu Pro Lys Glu Val His Val
          50          55          60
Pro Arg Asn Leu Asp Lys Glu Met Ile Gln Ser Val Val Asp Thr Lys
65          70          75          80

```

EP 0 841 394 A2

... Ile Val Gln Pro Ala Arg Gly Pro Lys Lys Asp Met Val Asp Leu Ala ...  
 5                               85                               90                               95  
 Ala His Asn Ala Lys Val Ser Leu Asn Asn Lys Phe Glu Leu Ile Ser  
                              100                               105                               110  
 Arg Asp Glu Ser Arg Thr Ile Lys Ala Ile Glu Glu Leu Gly Thr Gln  
                              115                               120                               125  
 10 Met Gly Ile Gln Thr Pro Ile Arg Ile Glu Ala Phe Asp Asn Ser Asn  
                              130                               135                               140  
 Ile Gln Gly Val Asp Pro Val Ser Ala Met Val Thr Phe Val Asp Gly  
                              145                               150                               155                               160  
 15 Lys Pro Asp Lys Lys Asn Tyr Arg Lys Tyr Lys Ile Lys Thr Val Lys  
                              165                               170                               175  
 Gly Pro Asp Asp Tyr Lys Ser Met Arg Glu Val Val Arg Arg Arg Tyr  
                              180                               185                               190  
 20 Ser Arg Val Leu Asn Glu Gly Leu Pro Leu Pro Asp Leu Ile Ile Val  
                              195                               200                               205  
 Asp Gly Gly Lys Gly His Met Asn Gly Val Ile Asp Val Leu Gln Asn  
                              210                               215                               220  
 25 Glu Leu Gly Leu Asp Ile Pro Val Ala Gly Leu Gln Lys Asn Asp Lys  
                              225                               230                               235                               240  
 His Gln Thr Ser Glu Leu Leu Leu Trp Arg  
                              245                               250  
 30

(2) INFORMATION FOR SEQ ID NO:469:

35                               (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 68 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 40                               (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

45                               (xi) SEQUENCE DESCRIPTION: SEQ ID NO:469:

Val Val Gln Xaa Thr Arg Phe His Asn Ile Cys Pro His Lys Gln Gly  
   1                               5                               10                               15  
 50 Pro Leu Ser Glu Gly Thr Val Ser Gly Glu Tyr Val Phe Cys Pro Leu  
                              20                               25                               30  
 His Asp Gln Lys Ile Asp Leu Asn Thr Gly Ile Val Gln Glu Pro Asp  
                              35                               40                               45  
 55 Glu Gly Cys Val Asp Val Tyr Glu Val Glu Val Thr Asp Gly Asn Val  
                              50                               55                               60

EP 0 841 394 A2

Tyr Ile Cys Leu

65

5 (2) INFORMATION FOR SEQ ID NO:470:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 95 amino acids

10 (B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:470:

20 Val His Val Leu Ala Phe Leu Thr Lys His His Ser Glu Lys Phe Asn  
1 5 10 15  
Ser Ser Ser Leu Ala Glu Leu Thr Cys Leu Xaa Pro Val Gln Leu Arg  
20 25 30  
25 Arg Val Thr Thr Gln Leu Val Asp Leu Xaa Met Ile Asp Thr Ile Arg  
35 40 45  
Gly Lys Asp Gly Gly Tyr Leu Ala Asn Asp Gln Ser Ala Asp Val Ser  
50 55 60  
30 Leu Ala Thr Leu Tyr Lys His Phe Val Leu Glu Lys Glu Gln His Thr  
65 70 75 80  
Arg Leu Phe Thr Trp Arg Arg Arg Gln Ser Leu Ser Asn Cys Ser  
85 90 95

35

(2) INFORMATION FOR SEQ ID NO:471:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 162 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

45

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:471:

50

Val Gly Phe Leu Asn His Met Leu Thr Leu Phe Thr Phe His Ser Gly  
1 5 10 15  
55 Leu Ser Leu Asn Ile Glu Ala Gln Gly Asp Ile Asp Val Asp Asp His  
20 25 30

**EP 0 841 394 A2**

[illegible]

(2) INFORMATION FOR SEQ ID NO:472:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 67 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:472:

```

Val Asn Thr Ala Met Gly Asn Leu Lys Gln Gly Ile Ala Asn Lys Asp
1          5          10          15
Thr Val Lys Ala Ser Glu Asn Tyr His Asp Ala Asp Val Asp Lys Gln
45          20          25          30
Thr Ala Tyr Thr Asn Ala Val Ser Gln Ala Glu Gly Ile Ile Asn Gln
          35          40          45
Thr Thr Asn Pro Thr Leu Asn Pro Asp Asp Ile Thr Pro Cys Ile Asn
50          50          55          60
Ser Ser Asp
65

```

(2) INFORMATION FOR SEQ ID NO:473:

# EP 0 841 394 A2

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 99 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:473:

```

Val Glu Ala Gln Lys Asp Val Ala Asn Val Leu Glu Asn Val Glu Gln
 1             5             10             15
Val Asp Ala Xaa Val Gly Ala Thr Asp Thr Ile Ala Leu Ala Ala Tyr
      20             25             30
Lys Tyr Tyr Ser Asp Lys Asn Asp Val Met Lys Pro His Gln Ile Tyr
      35             40             45
Gly Xaa Gly Gly Asp Pro Met Thr Gln Leu Val Ser Pro Ser Ile Lys
      50             55             60
Thr Ile His Tyr Asn Tyr Cys Glu Ala Gly Gln Cys Ala Arg Trp Gln
      65             70             75             80
Arg Tyr Asn Arg Cys Leu Lys Ser Lys Ile Cys His Ile Ala Ser Gln
      85             90             95
Xaa Asp Cys

```

## (2) INFORMATION FOR SEQ ID NO:474:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 89 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Protein

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:474:

```

Val Asp His Glu Val Phe Gln Gln Phe Gly Glu Ser Leu Pro Val Tyr
 1             5             10             15
Lys Pro Thr Leu Pro Pro Met Val Phe Gly Asn Arg Asp Lys Lys Ile
      20             25             30
Xaa Gly Gly Thr Asp Ala Leu Val Arg Tyr Leu Thr Pro His Gly
      35             40             45

```

EP 0 841 394 A2

... Xaa Trp Asn Ile His Ser Met Tyr Gln Asp Asn Lys His Met Leu Thr ...  
 50 55 60  
 5 Leu Phe Arg Gly Val His Arg Phe Gly Tyr Xaa Met Lys Met Leu Xaa  
 65 70 75 80  
 Asn Thr Ile Ser Lys Ile Met Ile Gly  
 85

(2) INFORMATION FOR SEQ ID NO:475:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 99 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:475:

Val Tyr Xaa Arg Asn Gly Val Val Thr Ala Arg Ala Val Ile Ser His  
 1 5 10 15  
 Arg Met Pro Lys Gly Thr Met Phe Met Tyr His Ala Gln Asp Lys His  
 20 25 30  
 Ile Gln Thr Pro Gly Ser Glu Ile Thr Asp Thr Arg Gly Gly Ser His  
 35 40 45  
 Asn Ala Pro Thr Arg Ile His Leu Lys Pro Thr Gln Leu Val Gly Gly  
 50 55 60  
 Tyr Ala Gln Ile Ser Tyr His Phe Asn Tyr Tyr Gly Pro Ile Gly Asn  
 65 70 75 80  
 Gln Arg Asp Leu Tyr Val Ala Val Arg Lys Met Lys Glu Val Asn Trp  
 85 90 95  
 Leu Glu Asp

(2) INFORMATION FOR SEQ ID NO:476:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 173 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:476:

5 Val Thr Trp Gly Thr Asn Pro Glu Met Gly Val Asn Phe Ser Glu Pro  
1 5 10 15  
Phe Pro Glu Ile Asn Asp Ile Asn Asp Gln Arg Ala Tyr Asp Tyr Met  
20 25 30  
10 Gly Leu Glu Pro Gly Gln Lys Ala Glu Asp Ile Asp Leu Gly Tyr Val  
35 40 45  
Phe Leu Gly Ser Cys Thr Asn Ala Arg Leu Ser Asp Leu Ile Glu Ala  
50 55 60  
15 Ser His Ile Val Lys Gly Asn Lys Val His Pro Asn Ile Thr Ala Ile  
65 70 75 80  
Val Val Pro Gly Ser Arg Thr Val Lys Lys Glu Ala Glu Lys Leu Gly  
85 90 95  
20 Leu Asp Thr Ile Phe Lys Asn Ala Gly Phe Glu Trp Arg Glu Pro Gly  
100 105 110  
Cys Ser Met Cys Leu Gly Met Asn Pro Asp Gln Val Pro Glu Gly Val  
115 120 125  
25 His Cys Ala Ser Thr Ser Asn Arg Asn Phe Glu Gly Arg Gln Gly Lys  
130 135 140  
Gly Ala Arg Thr His Leu Val Ser Pro Ala Met Ala Ala Ala Ala  
145 150 155 160  
30 Ile His Gly Lys Phe Val Asp Val Arg Lys Val Val Val  
165 170

(2) INFORMATION FOR SEQ ID NO:477:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 264 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:477:

50 Val Trp Asn Arg His Val Leu Xaa Gly Lys Xaa Gly Asp Pro Gln Leu  
1 5 10 15  
Leu Tyr Ile Asp Leu His Leu Ile His Glu Val Thr Ser Pro Gln Ala  
20 25 30  
55 Phe Glu Gly Leu Arg Leu Gln Asn Arg Lys Leu Arg Arg Pro Asp Leu  
35 40 45



EP 0 841 394 A2

... Thr Phe Ala Thr Leu Asp His Asn Val Pro Thr Ile Asp Ile Phe Asn ...  
 50 55 60  
 5 Ile Lys Asp Glu Ile Ala Asn Lys Gln Ile Thr Thr Leu Gln Lys Asn  
 65 70 75 80  
 Ala Ile Asp Phe Gly Val His Ile Phe Asp Met Gly Ser Asp Glu Gln  
 85 90 95  
 10 Gly Ile Val His Met Val Gly Pro Glu Thr Gly Leu Thr Gln Pro Gly  
 100 105 110  
 Lys Thr Ile Val Cys Gly Asp Ser His Thr Ala Thr His Gly Ala Phe  
 115 120 125  
 15 Gly Ala Ile Ala Phe Gly Ile Gly Thr Ser Glu Val Glu His Val Phe  
 130 135 140  
 Ala Thr Gln Thr Leu Trp Gln Thr Lys Pro Lys Asn Leu Lys Ile Asp  
 145 150 155 160  
 20 Ile Asn Gly Thr Leu Pro Thr Gly Val Tyr Ala Lys Asp Ile Ile Leu  
 165 170 175  
 His Leu Ile Lys Thr Tyr Gly Val Asp Phe Gly Thr Gly Tyr Ala Leu  
 180 185 190  
 25 Glu Phe Thr Gly Glu Thr Ile Lys Asn Leu Ser Met Asp Gly Arg Met  
 195 200 205  
 Thr Ile Cys Asn Met Ala Ile Glu Gly Gly Ala Lys Tyr Gly Ile Ile  
 210 215 220  
 30 Gln Pro Asp Asp Ile Thr Phe Glu Tyr Val Lys Gly Arg Pro Phe Ala  
 225 230 235 240  
 Asp Asn Phe Ala Lys Ser Val Asp Lys Trp Arg Glu Leu Tyr Ser Asp  
 245 250 255  
 35 Gly Thr Thr Arg Tyr Leu Ile Val  
 260

(2) INFORMATION FOR SEQ ID NO:478:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:478:

Val Ile Ile Asn Lys Val Met Thr Ser Asp Thr Pro Val Thr Ile Val  
 55 1 5 10 15

EP 0 841 394 A2

Ala Thr Gly Pro Leu Thr Asn Val Ala Thr Ala Leu Ile Arg Asp Pro  
 20 25 30  
 Arg Ile Ala Glu His Ile Glu Ser Ile Thr Leu Met Gly Gly Gly Thr  
 35 40 45  
 Phe Gly Asn Trp Thr Ala Tyr Ser Arg Ile Leu Ser Ile Pro Thr Ser  
 50 55 60  
 Phe Leu Thr Lys Ser Xaa Cys Gly Phe Val Asn Met Pro Leu Gly Val  
 65 70 75 80  
 Ile

(2) INFORMATION FOR SEQ ID NO:479:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 54 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:479:

Val Ala His Xaa His Val Val Asn Gly Thr Tyr Tyr Leu His Xaa His  
 1 5 10 15  
 Ile Val Xaa Gly Trp Gln Gly Val Lys Lys Thr Cys Asp Thr Ala Glu  
 20 25 30  
 Glu Leu Asp Thr Tyr Ile Lys Xaa Ser Asp Val Val Tyr Glu Glu Gln  
 35 40 45  
 Lys His Leu Xaa Leu Phe  
 50

(2) INFORMATION FOR SEQ ID NO:480:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:480:

EP 0 841 394 A2

... Val Ala His Ile His Val Val Asn Gly Thr Tyr Tyr Phe His Gly His ...  
 1 5 10 15  
 Ile Val Pro Gly Trp Gln Gly Val Lys Lys Thr Phe Asp Thr Ala Glu  
 20 25 30  
 Glu Leu Glu Thr Tyr Ile Lys Gln Ser Asp Ser Gly Ile  
 35 40 45

(2) INFORMATION FOR SEQ ID NO:481:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 148 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:481:

Val Cys Thr Tyr Cys Gly Val Gly Cys Ser Phe Glu Val Trp Thr Lys  
 1 5 10 15  
 Asp Arg Glu Ile Leu Lys Val Gln Pro Ser His Asp Ser Pro Ala Asn  
 20 25 30  
 Lys Ile Ala Thr Cys Val Lys Gly Lys Phe Ser Trp Gly His Ile Asn  
 35 40 45  
 Ser Asp Gln Arg Leu Thr Lys Pro Leu Val Arg Lys Asn Gly Glu Phe  
 50 55 60  
 His Glu Val Glu Trp Asp Glu Ala Leu Asn Val Ile Ala Asp Asn Phe  
 65 70 75 80  
 Thr Ser Ile Lys Glu Lys Tyr Gly Pro Asp Ala Leu Ser Phe Ile Ser  
 85 90 95  
 Ser Ser Lys Ala Thr Asn Glu Glu Ser Xaa Leu Xaa Gln Lys Leu Ala  
 100 105 110  
 Arg Gln Val Ile Gly Thr Asn Asn Val Asp Ser Leu Xaa Lys Asp Ile  
 115 120 125  
 Xaa Lys His Leu Gln Gln Lys Ala Tyr Leu Glu Arg Leu Asp Thr Ala  
 130 135 140  
 Gly Asp Ser Arg  
 145

(2) INFORMATION FOR SEQ ID NO:482:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 106 amino acids

EP 0 841 394 A2

(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:482:

```

Val Thr Lys Ala Val Gly Glu Arg Ile Pro Ile Thr Ile Ile Val Ala
 1           5           10           15
Val Leu Ala Leu Ile Val Ala Leu Ile Ile Ala Ile Pro Ile Gly Ile
 20           25           30
Ile Ser Ala Met Lys Arg Asn Ser Trp Leu Asp Ile Thr Leu Met Ile
 35           40           45
Ile Ala Leu Ile Gly Leu Ser Ile Pro Ser Phe Trp Gln Gly Leu Leu
 50           55           60
Phe Ile Leu Ala Xaa Ser Leu Lys Leu Asp Ile Leu Pro Pro Ser Tyr
 65           70           75           80
Met Pro Glu His Pro Ile Ser Val Asp Phe Thr Cys Thr Cys His Trp
 85           90           95
Asn Lys Tyr Cys Cys Phe Tyr His Ala Tyr
100           105

```

(2) INFORMATION FOR SEQ ID NO:483:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 46 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:483:

```

Val Ala Glu Leu Leu Gly Gly Ser Ala Val Thr Glu Gln Val Phe Asn
 1           5           10           15
Ile Asn Gly Ile Gly Arg Tyr Ile Val Gln Lys Gln Leu Ile Pro Asp
 20           25           30
Ile Pro Ala Val Met Gly Gly Val Val Tyr Ile Tyr Gln Leu
 35           40           45

```

(2) INFORMATION FOR SEQ ID NO:484:

EP 0 841 394 A2

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 152 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:484:

```

Val Thr His Ile Ile Ile Asp Gly Asp Ala Cys Pro Val Xaa Asp Ser
 1           5           10           15
Ile Ile Asp Leu Thr Thr Glu Thr Gly Ile Phe Val Thr Ile Ile Arg
 20           25           30
Ser Phe Ser His Phe Ser Asn Gln Leu Tyr Pro Pro His Val Ser Thr
 35           40           45
Leu Tyr Val Asp Asp Gly Pro Asp Ala Val Asp Tyr Lys Ile Val Gln
 50           55           60
Leu Ser Thr Lys Asp Asp Ile Val Val Thr Gln Asp Tyr Gly Leu Ala
 65           70           75           80
Ser Leu Leu Val Asp Lys Val Leu Ile Val Met His His Asn Gly Lys
 85           90           95
Ile Tyr Asn Ser Lys Asn Ile Gln Gln Leu Leu Asp Lys Arg Tyr Met
100          105          110
Asn Ala Gln Ile Arg Lys Gln Gly Gly Arg His Lys Gly Pro Pro Pro
115          120          125
Phe Thr Lys Gln Asp Gln Lys Val Phe Glu Gln Ser Leu Leu Lys Val
130          135          140
Ile His Arg Ile Lys Glu Leu Asp
145          150

```

(2) INFORMATION FOR SEQ ID NO:485:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 66 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:485:

EP 0 841 394 A2

Val Ile Glu Ser Leu Pro Glu His Val Asp Thr Ile Ile Asp Ile Lys  
 1 5 10 15  
 Ser Arg Thr Glu Gly Glu Leu Ile Thr Lys Glu Lys Glu Leu Val Gln  
 20 25 30  
 Leu Lys Phe Thr Pro Glu Asn Ile Asp Asn Val Asp Lys Glu Tyr His  
 35 40 45  
 Arg Glu Thr Phe Gly Glu Phe Asp Thr Arg Arg Thr Phe Glu Lys Cys  
 50 55 60  
 Asn Ser  
 65

(2) INFORMATION FOR SEQ ID NO:486:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 117 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:486:

Val Asp Ser Ile Ile Leu Asp Ala Ala Gln Thr Pro Leu Val Ile Ser  
 1 5 10 15  
 Gly Ala Pro Arg Leu Gln Ser Asn Leu Phe His Ile Val Lys Glu Phe  
 20 25 30  
 Val Asp Thr Leu Ile Glu Asp Val His Phe Lys Met Lys Lys Thr Lys  
 35 40 45  
 Lys Glu Ile Trp Leu Leu Asn Gln Gly Ile Glu Ala Ala Gln Ser Tyr  
 50 55 60  
 Phe Asn Val Xaa Asp Leu Tyr Ser Glu Gln Ala Met Val Leu Val Arg  
 65 70 75 80  
 Asn Ile Asn Leu Ala Leu Arg Ser Cys Arg Met Phe Leu Gly Ser Tyr  
 85 90 95  
 Lys Ser Tyr Arg Cys Lys Asn Val Leu Lys Phe Gly Cys Met Ile Gly  
 100 105 110  
 Glu Gln Pro Gly Gly  
 115

(2) INFORMATION FOR SEQ ID NO:487:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 270 amino acids

EP 0 841 394 A2

(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:487:

Val	Val	Ser	Leu	Ala	Xaa	Val	Leu	Ile	Ile	Gly	Ile	Gly	Ala	Leu	Thr
1				5					10					15	
Asn	Ser	Ile	Ser	Xaa	Asn	Leu	Gly	Arg	Asn	Thr	Val	Leu	Val	Leu	Gly
			20					25					30		
Ile	Ala	Leu	Pro	Xaa	Ile	Tyr	Phe	Thr	Leu	Ile	Ile	Arg	Ser	Xaa	Asp
			35				40					45			
Val	Xaa	Asp	Thr	Xaa	Arg	Ser	Arg	Val	Lys	Ala	Phe	Ile	Pro	Leu	Phe
		50				55				60					
Ile	Leu	Gly	Met	Val	Phe	Trp	Ala	Ile	Gln	Glu	Gln	Gly	Ser	Asn	Val
65					70				75					80	
Leu	Asn	Ile	Tyr	Gly	Xaa	Xaa	His	Ser	Asp	Met	Lys	Leu	Asn	Leu	Phe
				85					90				95		
Gly	Trp	Lys	Thr	Xaa	Phe	Gly	Glu	Ala	Ile	Phe	Gln	Ser	Ile	Asn	Pro
			100					105					110		
Leu	Phe	Ile	Leu	Leu	Leu	Ala	Pro	Ile	Ile	Ser	Leu	Leu	Trp	Gln	Lys
		115				120					125				
Xaa	Gly	Thr	Lys	Gln	Pro	Ser	Leu	Pro	Val	Lys	Phe	Ala	Ile	Gly	Thr
	130					135					140				
Phe	Leu	Ala	Gly	Ala	Ser	Tyr	Ile	Leu	Ile	Gly	Ile	Val	Gly	Tyr	Ala
145					150					155				160	
Ser	Gly	Ser	Ser	Asn	Phe	Ser	Val	Asn	Trp	Val	Ile	Leu	Ser	Tyr	Ile
				165					170				175		
Ile	Cys	Val	Ile	Gly	Glu	Leu	Cys	Leu	Ser	Pro	Thr	Gly	Asn	Ser	Ala
			180					185					190		
Ala	Val	Lys	Leu	Ala	Pro	Lys	Ala	Phe	Asn	Ala	Gln	Met	Met	Ser	Ile
		195				200					205				
Trp	Tyr	Leu	Thr	Asn	Ala	Ser	Ala	Gln	Ala	Ile	Asn	Gly	Thr	Leu	Val
	210					215					220				
Lys	Leu	Ile	Glu	Pro	Leu	Gly	Gln	Thr	Asn	Tyr	Phe	Ile	Phe	Leu	Gly
225					230					235				240	
Val	Val	Ala	Ile	Ile	Val	Thr	Thr	Ile	Cys	Ile	Ser	Ile	Leu	Thr	Phe
				245					250				255		
Asn	His	Gln	Ser	Asp	Glu	Arg	Tyr	Thr	Leu	Ile	Leu	Leu	Ala		
				260					265				270		

(2) INFORMATION FOR SEQ ID NO:488:

EP 0 841 394 A2

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 80 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:488:

```

Val Thr Pro Pro His Pro Asn Gly Val Ser Gln Glu Val Leu Ala Ala
 1           5           10           15
Cys Tyr Leu Thr Gln Xaa Thr Gln Val Xaa Xaa Xaa Gly Gly Ala Xaa
          20           25           30
Ser Ile Ala Xaa Leu Thr Tyr Xaa Ala Xaa Thr Ile Pro Lys Val Asp
          35           40           45
Lys Ile Xaa Gly Pro Gly Thr Gln Phe Val Ala Ser Xaa Xaa Lys Xaa
          50           55           60
Leu Phe Gly Gln Val Ala Tyr Cys Ser His Xaa Arg Xaa Thr Xaa Gln
65           70           75           80

```

(2) INFORMATION FOR SEQ ID NO:489:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:489:

```

Val Phe Gln Val Gly Gly Ala Gln Ser Ile Ala Ala Leu Thr Tyr Gly
 1           5           10           15
Thr Glu Thr Ile Pro Lys Val Asp Lys Ile Val Gly Pro Gly Asn Gln
          20           25           30
Phe Gly Xaa Ile Cys Pro Lys Asn Ile Tyr Leu Asp Arg
          35           40           45

```

(2) INFORMATION FOR SEQ ID NO:490:

(i) SEQUENCE CHARACTERISTICS:



EP 0 841 394 A2

(A) LENGTH: 96 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:490:

Val	Gly	Met	Ser	Ile	Val	Leu	Ala	Gly	Gly	Met	Glu	Asn	Met	Ser	Gln
1				5				10					15		
Xaa	Pro	Met	Leu	Xaa	Asn	Asn	Ser	Arg	Phe	Gly	Phe	Lys	Met	Gly	His
			20					25					30		
His	Ser	Met	Val	Asp	Ser	Met	Val	Tyr	Asp	Gly	Leu	Thr	Asp	Val	Phe
			35					40					45		
Thr	Gln	Tyr	His	Met	Gly	Ile	Thr	Ala	Glu	Asn	Leu	Val	Gly	Ala	Ile
			50					55					60		
Trp	Tyr	Phe	Lys	Arg	Arg	Thr	Lys	Ile	His	Val	Ala	Gly	Asn	Leu	Thr
			65					70					75		80
Thr	Ile	Lys	Gln	Tyr	Val	His	Ser	Lys	Met	Val	Asn	Leu	Ile	Val	Lys
			85					90					95		

(2) INFORMATION FOR SEQ ID NO:491:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 130 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:491:

Val	Ile	Ser	Met	Lys	Trp	Leu	Ser	Arg	Ile	Leu	Thr	Val	Ile	Val	Thr
1				5					10				15		
Met	Ser	Met	Ala	Cys	Gly	Ala	Leu	Ile	Phe	Asn	Arg	Arg	His	Gln	Leu
			20						25				30		
Lys	Thr	Lys	Thr	Leu	Asn	Phe	Asn	His	Lys	Ala	Leu	Thr	Ile	Ile	Ile
			35						40				45		
Pro	Ala	Arg	Asn	Glu	Glu	Lys	Arg	Ile	Gly	His	Leu	Leu	His	Ser	Ile
			50					55					60		
Ile	Gln	Gln	Gln	Val	Pro	Val	Asp	Val	Ile	Val	Met	Asn	Asp	Gly	Ser
			65					70					75		80

EP 0 841 394 A2

.. Thr Asp Glu Thr Ala Arg Val Ala Arg Ser Tyr Gly Ala Thr Val Val w .  
 85 90 95  
 Asp Val Val Asp Asp Thr Asp Gly Lys Trp Tyr Gly Glu Ile Thr Cys  
 100 105 110  
 Leu Leu Ser Arg Cys Asp Ala Cys Met Tyr Glu Ser Tyr Cys Leu Cys  
 115 120 125  
 Arg Cys  
 130

(2) INFORMATION FOR SEQ ID NO:492:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 124 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:492:

Val Pro Ile Leu Gly Thr Val Phe Thr Ala Leu Thr Gly Pro Ile Gly  
 1 5 10 15  
 Ile Val Leu Gly Val Leu Ala Gly Leu Ala Val Ala Phe Thr Ile Ala  
 20 25 30  
 Tyr Lys Lys Ser Glu Thr Phe Arg Asn Cys Val Asn Gly Ala Ile Asn  
 35 40 45  
 Ser Val Lys Gln Thr Phe Ser Asn Phe Ile Gln Phe Ile Gln Pro Tyr  
 50 55 60  
 Ile Asp Ser Val Lys Asn Val Phe Lys Gln Ala Val Ser Ala Ile Gly  
 65 70 75 80  
 Asp Phe Ala Lys Asp Ile Trp Ser Gln Ile Asn Gly Phe Phe Asn Glu  
 85 90 95  
 Asn Gly Ile Ser Ile Ala Gln Ala Leu Gln Asn Ile Cys Asn Phe Ile  
 100 105 110  
 Lys Ala Ile Phe Glu Phe Ile Leu Lys Phe Cys Asn  
 115 120

(2) INFORMATION FOR SEQ ID NO:493:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 147 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single

# EP 0 841 394 A2

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:493:

```

Val Lys Ser Leu Lys Thr Val Ile Gly Met Asn Asn Lys Glu His Ile
 1           5           10           15
Lys Ser Val Ile Leu Ala Leu Leu Val Leu Met Ser Val Val Leu Thr
          20           25           30
Tyr Met Val Trp Asn Phe Ser Pro Asp Ile Ala Asn Val Asp Asn Thr
          35           40           45
Asp Ser Lys Lys Ser Glu Thr Lys Pro Leu Thr Thr Pro Met Thr Ala
          50           55           60
Lys Met Asp Thr Thr Ile Thr Pro Phe Gln Ile Ile His Ser Lys Asn
        65           70           75           80
Asp His Pro Glu Gly Thr Ile Ala Thr Val Ser Asn Val Asn Lys Leu
          85           90           95
Thr Lys Pro Leu Lys Asn Lys Glu Val Xaa Ser Val Glu His Val Arg
          100          105          110
Arg Asp His Asn Leu Met Ile Pro Asp Leu Ser Ser Asp Phe Thr Leu
          115          120          125
Phe Asp Phe Thr Phe Asp Leu Pro Leu Ser Thr Tyr Leu Gly Gln Val
          130          135          140
Xaa Glu His
145

```

(2) INFORMATION FOR SEQ ID NO:494:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 90 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:494:

```

Val Ile Leu Pro Lys Leu Val Thr Tyr Gln Val Gln Pro Met Ser Met
 1           5           10           15
Met Leu Ile Pro Lys Gly Xaa Xaa Glu Leu Ala Leu Arg Phe Asn Gln
          20           25           30

```

# EP 0 841 394 A2

... Gly Ile Asn Xaa Gly Gly Asp Ala Leu Lys Lys Arg Thr His Phe Asn ...  
 35 40 45  
 5 Ile Ala Gly Ala Phe Asn Pro Asp Val Arg Lys Leu Asp Gly Ala Val  
 50 55 60  
 Xaa Arg Leu Glu Xaa Lys Asp Xaa Lys Arg Asn Val Leu Phe Xaa Asn  
 65 70 75 80  
 10 Thr Thr Arg Val Gln Gln Arg Glu Asn His  
 85 90

## (2) INFORMATION FOR SEQ ID NO:495:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 126 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Protein

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:495:

Val Leu Val Gln Pro His Ile Val Ile Lys Pro Glu Ala Gln Gln Ala  
 1 5 10 15  
 30 Ile Lys Ala Thr Ala Glu Asn Gln Val Glu Ser Ile Lys Asp Thr Pro  
 20 25 30  
 His Ala Thr Val Asp Glu Leu Asp Glu Ala Asn Gln Leu Ile Ser Asp  
 35 35 40 45  
 Thr Leu Lys Gln Ala Gln Gln Glu Ile Glu Asn Thr Asn Gln Asp Ala  
 50 55 60  
 Ala Val Thr Asp Val Arg Asn Gln Thr Ile Lys Ala Ile Glu Gln Ile  
 40 65 70 75 80  
 Lys Pro Lys Val Arg Arg Lys Arg Ala Ala Leu Asp Ser Ile Glu Glu  
 85 90 95  
 Asn Xaa Lys Asn Gln Leu Gly Cys Asn Pro Glu Ile Arg Gly Ile Leu  
 45 100 105 110  
 Leu Lys Asp Glu Arg Asp Cys Trp Leu Leu Asp Thr Phe Lys  
 115 120 125

## (2) INFORMATION FOR SEQ ID NO:496:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single

EP 0 841 394 A2

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:496:

Val Ser Leu Leu Pro Leu Ile Ser Asp Asn Met Tyr Leu Gly Tyr Ile  
1 5 10 15  
Ala Gly Leu Leu Ala Ala Tyr Ala Gly Gly Phe Val Cys Thr Tyr Leu  
20 25 30  
Phe Gly Pro Thr Lys Ala Met Arg Gln Asp Arg Phe Val Gly  
35 40 45

(2) INFORMATION FOR SEQ ID NO:497:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 90 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:497:

Val Gly Xaa Xaa Ile Met Lys Phe Lys Ala Ile Ala Lys Xaa Ser Leu  
1 5 10 15  
Ala Leu Gly Met Leu Ala Thr Gly Val Ile Thr Ser Asn Val Gln Ser  
20 25 30  
Val Gln Ala Xaa Thr Glu Val Xaa Gln Gln Ser Glu Ser Glu Leu Xaa  
35 40 45  
His Tyr Tyr Asn Lys Pro Val Leu Glu Arg Lys Asn Val Thr Gly Tyr  
50 55 60  
Xaa Tyr Thr Glu Lys Gly Lys Asp Tyr Ile Asp Val Ile Val Asp Asn  
65 70 75 80  
Gln Tyr Ser Gln Ile Ser Leu Val Arg Ile  
85 90

(2) INFORMATION FOR SEQ ID NO:498:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 147 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single

EP 0 841 394 A2

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:498:

Val Ile Asn Ile Glu Leu Val Leu His Gln Arg Xaa Ala Asn Pro Val  
 1 5 10 15  
 Arg Ile Ala Asn Ile Ser Asn Asn Ala Thr Val Ser Gln Ala Asp Gln  
 20 25 30  
 Ala Thr Ile Ile Asn Ser Leu Thr Phe Thr Ser Asn Ala Pro Asn Arg  
 35 40 45  
 Asn Tyr Ala Thr Ala Ser Ala Asn Glu Ile Thr Ser Lys Thr Val Ser  
 50 55 60  
 Asn Val Ser Arg Thr Gly Asn Asn Ala Asn Val Thr Val Thr Val Thr  
 65 70 75 80  
 His Gln Asp Gly Thr Thr Ser Thr Val Thr Val Pro Val Lys His Val  
 85 90 95  
 Ile Pro Glu Ile Val Ala His Ser His Tyr Thr Val Gln Gly Gln Asp  
 100 105 110  
 Phe Pro Ala Gly Asn Gly Ser Ser Ala Ala Asp Tyr Phe Lys Leu Ser  
 115 120 125  
 Asn Gly Ser Ala Ile Pro Asp Ala Thr Ile Thr Trp Val Ser Gly Gln  
 130 135 140  
 Xaa Pro Lys  
 145

(2) INFORMATION FOR SEQ ID NO:499:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 274 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:499:

Val Met Glu Asn Ser Arg Pro Glu Arg Asn Glu Ala Thr Met His Leu  
 1 5 10 15  
 Asp Glu Met Thr Val Glu Glu Ala Leu Ile Thr Met Asn Lys Glu Asp  
 20 25 30

EP 0 841 394 A2

... Gln Gln Val Pro Leu Ala Val Arg Lys Ala Ile Pro Gln Leu Thr Lys w . . .  
 35 40 45  
 Val Ile Lys Lys Thr Ile Ala Gln Tyr Lys Lys Gly Gly Arg Leu Ile  
 5 50 55 60  
 Tyr Ile Gly Ala Gly Thr Ser Gly Arg Leu Gly Val Leu Asp Ala Ala  
 65 70 75 80  
 Glu Cys Val Pro Thr Phe Asn Thr Asp Pro His Glu Ile Ile Gly Ile  
 10 85 90 95  
 Ile Ala Gly Gly Gln His Ala Met Thr Met Ala Val Glu Gly Ala Glu  
 100 105 110  
 Asp His Lys Lys Leu Ala Glu Glu Asp Leu Lys Asn Ile Asp Leu Thr  
 15 115 120 125  
 Ser Lys Asp Val Val Ile Gly Ile Ala Ala Ser Gly Lys Thr Pro Tyr  
 130 135 140  
 Val Ile Gly Gly Leu Thr Phe Ala Asn Thr Ile Gly Ala Thr Thr Val  
 20 145 150 155 160  
 Ser Ile Ser Cys Asn Glu His Ala Val Ile Ser Glu Ile Ala Gln Tyr  
 165 170 175  
 Pro Val Glu Val Lys Val Gly Pro Glu Val Leu Thr Gly Ser Xaa Arg  
 25 180 185 190  
 Leu Lys Ser Gly Thr Ala Gln Lys Leu Ile Leu Asn Met Ile Ser Thr  
 195 200 205  
 Ile Thr Met Val Gly Val Gly Lys Val Tyr Asp Asn Leu Met Ile Asp  
 30 210 215 220  
 Val Lys Ala Thr Asn Gln Lys Leu Ile Asp Arg Ser Val Arg Ile Ile  
 225 230 235 240  
 Gln Glu Ile Cys Ala Ile Thr Tyr Asp Glu Ala Met Ala Leu Tyr Gln  
 35 245 250 255  
 Val Ser Glu His Asp Val Glu Ser Cys Asp Ser Tyr Gly Tyr Val Trp  
 260 265 270  
 His Phe

(2) INFORMATION FOR SEQ ID NO:500:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 228 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:500:

EP 0 841 394 A2

Val Xaa Xaa Ile His Ser Gln Gln Ile Val Ile Ala Pro Asp Ser Phe  
 1 5 10 15  
 Lys Glu Ser Met Xaa Xaa His Gln Val Xaa Asn Ile Ile Lys Gln Ala  
 5 20 25 30  
 Phe Thr Asn Val Tyr Gly Asn Thr Leu His Tyr Asp Ile Ile Pro Met  
 35 40 45  
 Pro Asp Gly Gly Glu Gly Thr Xaa Asp Xaa Leu Ile His Ala Xaa Xaa  
 10 50 55 60  
 Ala Thr Lys Tyr Thr Val Ile Val Asn Asp Pro Leu Met Arg Pro Ile  
 65 70 75 80  
 Glu Ala Cys Tyr Ala Arg Ala Asp Glu Gln Gln Ile Ala Ile Ile Glu  
 15 85 90 95  
 Met Ala Ala Ala Ser Gly Leu Asp Leu Leu Glu Lys Glu Glu Arg Asn  
 100 105 110  
 Pro Leu Tyr Thr Ser Ser Tyr Gly Thr Gly Glu Leu Ile Lys Asp Ala  
 20 115 120 125  
 Leu Asn His Gly Ala Lys Thr Ile Ile Leu Gly Ile Gly Gly Ser Ala  
 130 135 140  
 Thr Asn Asp Gly Gly Thr Gly Met Leu Ser Ala Leu Gly Val Lys Phe  
 25 145 150 155 160  
 Thr Asp Val Asn Gly Asp Leu Leu Gln Met Asn Gly Ala Asn Leu Ala  
 165 170 175  
 His Ile Ala Gln Ile Asp Ile Thr Asn Leu Ala Asp Ser Arg Xaa Lys  
 30 180 185 190  
 Glu Val Thr Phe Lys Val Ala Cys Asp Val Ser Asn Pro Leu Leu Gly  
 195 200 205  
 Glu Asn Gly Ala Thr Tyr Ile Tyr Gly Pro Gln Lys Cys Ala Asp Ala  
 35 210 215 220  
 Lys Met His Thr  
 225

(2) INFORMATION FOR SEQ ID NO:501:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 117 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:501:



# EP 0 841 394 A2

1 Val Lys Gln Cys Ile Asn Asn Asp Glu Trp Phe Val Thr Asn Asp Asn  
 5 Gly Tyr Val Lys Glu Gln Tyr Leu Tyr Xaa Ala Gly Arg Gln Gln Asp  
 20 25 30  
 Met Leu Ile Ile Gly Gly Arg Asn Ile Tyr Pro Ala His Val Xaa Arg  
 35 40 45  
 10 Leu Leu Thr Gln Ser Ser Ser Ile Asp Glu Ala Ile Ile Ile Gly Ile  
 50 55 60  
 Pro His Glu Arg Phe Gly Xaa Ile Gly Val Leu Leu Tyr Ser Gly Asp  
 65 70 75 80  
 15 Val Thr Leu Thr His Lys Asn Xaa Lys Gln Phe Xaa Lys Lys Lys Val  
 85 90 95  
 Lys Arg His Glu Ile Pro Phe Asp Asp Ser Ser Cys Arg Lys Asp Val  
 100 105 110  
 20 Leu Xaa Cys Lys Trp  
 115

## (2) INFORMATION FOR SEQ ID NO:502:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 138 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Protein

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:502:

40 Val Ser Ala Ile Phe Glu Pro Glu His Leu Glu Ala Leu Leu Ala Gln  
 1 5 10 15  
 Val Ser Ile Lys Pro Val Ile Asn Gln Val Glu Tyr His Pro Tyr Leu  
 20 25 30  
 Thr Gln His Lys Leu Lys Leu Tyr Leu Ala Ala Gln His Ile Val Met  
 35 40 45  
 45 Glu Ser Trp Ser Pro Leu Met Asn Ala Gln Ile Leu Asn Asp Glu Thr  
 50 55 60  
 Ile Lys Asp Ile Ala Gln Glu Leu Gly Lys Ser Pro Ala Gln Val Val  
 50 65 70 75 80  
 Leu Arg Trp Asn Val Gln His Gly Val Val Ile Ile Pro Lys Ser Val  
 85 90 95  
 55 Thr Pro Asn Arg Ile Ser Glu Asn Phe Gln Ile Phe Asp Phe Glu Leu  
 100 105 110

EP 0 841 394 A2

... Ser Asp Glu Gln Met Thr Leu Val Pro Gly Leu Asn Leu Asp Lys Arg ...  
 115 120 125  
 5 Ile Gly Pro Asp Xaa Xaa Thr Phe Glu Gly  
 130 135

(2) INFORMATION FOR SEQ ID NO:503:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 78 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:503:

Val Phe Lys Gly Ser Glu Tyr Gln Leu Ser Glu Ile Asn Ser Gly Ser  
 1 5 10 15  
 Val Lys Tyr Glu Gln Thr Tyr Asp Asn Phe Pro Ile Leu Asn Asn Ser  
 20 25 30  
 Lys Ala Met Leu Asn Phe Asn Ile Glu Asp Asn Lys Ala Ala Ser Tyr  
 35 40 45  
 Lys Gln Ser Met Met Asp Asp Ile Lys Pro Thr Asp Gly Ala Asp Lys  
 50 55 60  
 Lys His Gln Val Ile Gly Val Arg Lys Ala Ile Glu Ala Leu  
 65 70 75

(2) INFORMATION FOR SEQ ID NO:504:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:504:

Val Asn Glu Thr Asn Val Gln Leu Leu Gln Pro Asn Trp Glu Ile Lys  
 1 5 10 15  
 Val Lys His Asp Gly Lys Asp Lys Thr Asn Thr Tyr Tyr Val Glu Ala  
 20 25 30

# EP 0 841 394 A2

Thr Asn Asn Asn Pro Lys Ile Ile Asn His...

35 40

## (2) INFORMATION FOR SEQ ID NO:505:

### (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Protein

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:505:

Val Leu Ile Tyr Phe Ile Glu Gly Leu Met Ser Xaa Asn Gln Pro Ala  
1 5 10 15  
Glu Ala Leu Glu Xaa Leu Ser Tyr Val Asp Pro Ser Pro Xaa Xaa Leu  
20 25 30  
Met Xaa Lys Lys  
35

## (2) INFORMATION FOR SEQ ID NO:506:

### (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 75 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Protein

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:506:

Val Gly Ile Asn Ser Lys Gln Gly Lys Asn Leu Ile Gly Ala Phe Tyr  
1 5 10 15  
Arg Pro Thr Ala Val Ile Tyr Asp Leu Asp Phe Leu Lys Thr Leu Pro  
20 25 30  
Phe Glu Gln Ile Leu Ser Gly Tyr Ala Glu Val Tyr Lys His Ala Leu  
35 40 45  
Leu Asn Gly Glu Ser Thr Thr Gln Glu Ile Glu Gln His Phe Lys Asp  
50 55 60  
Arg Glu Ile Leu Gln Pro Leu Asn Gly Ile Gly  
65 70 75

EP 0 841 394 A2

(2) INFORMATION FOR SEQ ID NO:507:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 153 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:507:

```

Val Val Thr Leu Phe Leu Cys Pro Xaa Asn Ser Phe Leu Met Ile His
  1             5             10             15
Asn Ser Trp Val Met Thr Val Gly Asn Ala Glu Glu Leu Arg Lys Thr
  20             25             30
Ala Asp Leu Leu Glu Lys Thr Asp Ala Val Ser Asn Ser Ala Tyr Leu
  35             40             45
Asp Lys Xaa Xaa Asp Leu Asp Gln Glu His Leu Lys Gln Met Leu Asp
  50             55             60
Ala Glu Thr Trp Leu Thr Ala Glu Glu Ala Leu Ser Phe Gly Leu Ile
  65             70             75             80
Asp Glu Ile Leu Gly Ala Asn Glu Ile Ala Ala Ser Ile Ser Lys Glu
  85             90             95
Gln Cys Arg Arg Xaa Glu Xaa Val Pro Glu Asp Leu Lys Lys Asp Val
  100            105            110
Xaa Lys Ile Thr Lys Ile Asp Asp Xaa Arg Tyr Asp Leu Asp Trp Leu
  115            120            125
Lys Pro Pro Lys Glu Ser Met Ser Leu Glu Glu Xaa Xaa Xaa Arg Xaa
  130            135            140
Xaa Ile Val Arg Ser Asn Ala Lys Phe
  145            150

```

(2) INFORMATION FOR SEQ ID NO:508:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 115 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

EP 0 841 394 A2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:508:

```

5   Val Lys Glu Ile Pro Asp Ala Ser Ile Ser Phe Glu Val Phe Ala Asp
    1           5           10           15
    Asp Leu Glu Thr Met Glu Lys Glu Ala Ala Ile Leu Lys Gln Tyr Gly
      20           25           30
10  Glu Asn Val Phe Val Lys Ile Pro Ile Val Asn Thr Lys Gly Glu Ser
    35           40           45
    Thr Ile Pro Leu Ile Lys Lys Leu Ser Ala Asp Asn Val Arg Leu Asn
    50           55           60
15  Val Xaa Ala Val Tyr Thr Ile Glu Gln Gly Lys Glu Ile Thr Glu Ala
    65           70           75           80
    Val Thr Glu Gly Val Pro Asn Ile Cys Phe Ser Ile Cys Arg Thr Tyr
      85           90           95
20  Cys Arg Tyr Arg Arg Arg Ser Ile Thr Ile Asn Glu Arg Gly Cys Lys
    100          105          110
    Ser Tyr Ala
      115

```

(2) INFORMATION FOR SEQ ID NO:509:

(i) SEQUENCE CHARACTERISTICS:

```

30  (A) LENGTH: 197 amino acids
    (B) TYPE: amino acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

```

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:509:

```

40  Val Tyr Xaa Arg Val Asn Glu Met Asn Ala Lys Glu Xaa Leu Val Asp
    1           5           10           15
    Xaa Leu Met Lys Thr Ser Ser Gln Leu Phe Lys Xaa His Gly Glu Val
    20           25           30
45  Xaa Met Gln Leu Xaa Leu Asn Asp Glu Leu Lys Leu Pro Ser Ile Xaa
    35           40           45
    Glu Ile Cys Val Glu Arg Lys Arg Leu Ser Asp Ile Val Lys Val Ile
    50           55           60
    Pro Gln Ser Tyr Ala Leu Leu Tyr Ile Asp Lys Gln Asp Gln Ala Arg
    65           70           75           80
55  Ala Lys Xaa Xaa Leu Ser Leu Xaa Lys Ile Ala Lys Val Tyr Val Gln
    85           90           95

```

EP 0 841 394 A2

... Tyr Asp Asp Thr Thr Ile Met Ser Ile Phe Val Tyr Asp Val Val Asn ...  
 100 105 110  
 5 Asp Glu Trp Ile Leu Arg Leu Asp Pro Asn Ile Arg Ile Pro Lys Ser  
 115 120 125  
 Asn Ile Tyr Phe His Ser Leu Asn Trp Asp Val Asp Tyr Ile Lys Pro  
 130 135 140  
 10 Glu Ile Val Leu Met Tyr Asp Leu Met Gln His His Gln Tyr His His  
 145 150 155 160  
 Tyr Ser Asn Tyr Lys Arg Val Ile Asp Xaa Leu Ser Tyr Tyr Gln Phe  
 165 170 175  
 15 Phe Ile Leu Lys Phe Val Val Gly Glu Xaa Arg Ile Lys Asp Ala Ile  
 180 185 190  
 Gln Lys Asn Asn Lys  
 195

(2) INFORMATION FOR SEQ ID NO:510:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 56 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:510:

Val Pro Pro Gln Gln Gln Ala Pro Thr Lys Gln Arg Pro Ala Lys Xaa  
 1 5 10 15  
 Glu Asn Asp Asp Lys Ala Ser Xaa Asp Glu Ser Lys Asp Lys Asp Asp  
 20 25 30  
 Asn Ala Ser Gln Asp Xaa Ser Asp Asp Thr Gln Lys Lys Thr Asp Asp  
 35 40 45  
 Asn Thr His Pro Ala Ala Arg Xaa  
 50 55

(2) INFORMATION FOR SEQ ID NO:511:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 77 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

EP 0 841 394 A2

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:511:

```

Val Leu Met Leu Lys Met Xaa Leu Ile Lys Lys Leu Thr Gln Met Glu
 1             5             10             15
Asp Leu Val Asn Gln Asn Asp Xaa Leu Thr Asp Glu Glu Lys Gln Xaa
10             20             25             30
Ala Ile Gln Val Ile Glu Glu His Lys Asn Glu Ile Ile Gly Asn Ile
35             40             45
Gly Asp Gln Thr Thr Asp Asp Gly Val Thr Arg Ile Thr Arg Ser Arg
15             50             55             60
Tyr Thr Asp Leu Lys Trp Gly Tyr Cys Asn Thr Gly Cys
65             70             75

```

(2) INFORMATION FOR SEQ ID NO:512:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 166 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:512:

```

Val Lys Trp Ile Lys Arg Leu Leu Lys Asp Leu Lys Lys Leu Ser Met
 1             5             10             15
Leu Phe Lys Asp Ile Leu Gln Leu Val Pro Lys Gln Leu Phe Gly Gly
40             20             25             30
Gly Asp Ala Glu Leu Gln Leu Thr Glu Ala Asp Tyr Leu Thr Ala Gly
35             40             45
Ile Asp Ile Val Val Gln Pro Pro Gly Lys Lys Leu Gln His Leu Ser
45             50             55             60
Leu Leu Ser Gly Gly Glu Arg Ala Leu Thr Ala Ile Ala Leu Leu Phe
65             70             75             80
Ala Ile Leu Lys Val Arg Ser Ala Pro Phe Val Ile Leu Asp Glu Val
50             85             90             95
Glu Ala Ala Leu Asp Glu Ala Asn Val Ile Arg Tyr Ala Lys Tyr Leu
100            105            110
Asn Glu Leu Ser Asp Glu Thr Gln Phe Ile Val Ile Thr His Arg Lys
55            115            120            125

```

EP 0 841 394 A2

... Gly Thr Met Glu Phe Ala Asp Arg Leu Tyr Gly Val Thr Met Gln Glu ...  
 130 135 140  
 5 Ser Gly Val Thr Lys Leu Val Ser Val Asn Leu Asn Thr Ile Asp Asp  
 145 150 155 160  
 Val Leu Lys Glu Glu Gln  
 165

(2) INFORMATION FOR SEQ ID NO:513:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 88 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:513:

Val Arg Asp Gln Phe Lys Phe Tyr His Val Arg His Glu Glu Val Ala  
 1 5 10 15  
 Ser Leu Ala Ala Ala Gly Tyr Thr Lys Leu Thr Gly Xaa Ile Gly Val  
 20 25 30  
 Ala Leu Ser Ile Gly Xaa Pro Gly Leu Ile His Leu Leu Asn Gly Met  
 35 40 45  
 Tyr Asp Ala Lys Met Asp Asn Val Leu Val Pro Ile Asn Ile Ile Trp  
 50 55 60  
 Thr Asn Xaa Ile Val Gln His Leu Glu Arg Lys His Phe Lys Lys Gln  
 65 70 75 80  
 Ile Tyr Lys Asn Tyr Val Lys Met  
 85

(2) INFORMATION FOR SEQ ID NO:514:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 153 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:514:



EP 0 841 394 A2

... Val Lys Leu Ser Xaa Tyr Tyr Pro Gln Gly Leu Arg Ser Leu Asn Gly ...  
 1 5 10 15  
 5 Gly Arg Met Ala Arg Phe Gly Arg Thr Pro Leu Leu Asp Ala Met Glu  
 20 25 30  
 Met Ala Asn Glu His Ile Met Val Ile Ala Met Ile Glu Asp Val Xaa  
 35 40 45  
 10 Gly Val Met Ala Ile Asp Asp Ile Ala Gln Val Glu Gly Leu Asp Met  
 50 55 60  
 Ile Val Glu Gly Ala Ala Gly Phe Ile Ala Val Thr Trp His Thr Xaa  
 65 70 75 80  
 15 Ala Asn Glu Arg Asp Asp Gln Val Thr Ser His Xaa Gln His Ile Xaa  
 85 90 95  
 Xaa Val Val Asn Ala His Gly Lys His Xaa Cys Ala Leu Pro Arg Glu  
 100 105 110  
 20 Asp Glu Asp Ile Ala Lys Trp Gln Ala Gln Gly Val Gln Thr Phe Ile  
 115 120 125  
 Leu Gly Thr Ser Gly Lys Ile Tyr Arg His Leu Ser Ala Ser Leu Ala  
 130 135 140  
 25 Thr Ser Lys Gln Lys Gly Asp Asp Gly  
 145 150

(2) INFORMATION FOR SEQ ID NO:515:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 105 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:515:

Val Leu Thr Ser Leu Asp Ala Asp Glu Ala Lys Arg Asn Ala Tyr Thr  
 1 5 10 15  
 Asn Ala Val Thr Gln Ala Xaa Gln Ile Leu Asn Lys Ala Gln Gly Pro  
 20 25 30  
 Asn Thr Ala Lys Asp Gly Val Glu Thr Ala Leu Gln Asn Val Gln Arg  
 35 40 45  
 Ala Lys Asn Glu Leu Ser Gly Asn Gln Asn Val Ala Asn Ala Lys Thr  
 50 55 60  
 55 Thr Ala Lys Asn Ala Leu Asn Asn Leu Thr Ser Ile Asn Asn Ala His  
 65 70 75 80

EP 0 841 394 A2

Lys Ala Ala Leu Lys Ser Gln Ile Glu Gly Ala Thr Thr Val Ala Gly  
 85 90 95  
 Val Asn Gln Val Ser Thr Met Ala Ser  
 100 105

(2) INFORMATION FOR SEQ ID NO:516:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 276 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:516:

Val Asn Thr Ala Lys Thr Ala Leu Asn Gly Asp Ala Arg Leu Asn Glu  
 1 5 10 15  
 Ala Lys Asn Thr Ala Lys Gln Gln Leu Ala Thr Met Ser His Leu Thr  
 20 25 30  
 Asn Ala Gln Lys Ala Asn Leu Thr Glu Gln Ile Glu Arg Gly Thr Thr  
 35 40 45  
 Val Ala Gly Val Gln Gly Ile Gln Ala Asn Ala Gly Thr Leu Asn Gln  
 50 55 60  
 Ala Met Asn Gln Leu Arg Gln Ser Ile Ala Ser Lys Asp Ala Thr Lys  
 65 70 75 80  
 Ser Ser Glu Asp Tyr Gln Asp Ala Asn Ala Asp Leu Gln Asn Ala Tyr  
 85 90 95  
 Asn Asp Ala Val Thr Asn Ala Gly Gly Ile Ile Ser Ala Xaa Asn Asn  
 100 105 110  
 Pro Glu Met Asn Pro Asp Thr Ile Xaa Gln Lys Ala Ser Gln Val Asn  
 115 120 125  
 Ser Ala Lys Ser Ala Leu Xaa Gly Asp Glu Lys Leu Ala Ala Ala Lys  
 130 135 140  
 Gln Thr Ala Lys Ser Asp Ile Gly Arg Val Thr Asp Leu Asn Asn Ala  
 145 150 155 160  
 Gln Arg Thr Ala Xaa Asn Ala Glu Val Asp Gln Ala Pro Xaa Leu Ala  
 165 170 175  
 Ala Val Thr Ala Ala Lys Asn Lys Ala Thr Ser Leu Asn Thr Ala Met  
 180 185 190  
 Gly Asn Val Lys His Ala Leu Ala Glu Lys Asp Asn Thr Xaa Arg Ser  
 195 200 205

EP 0 841 394 A2

Val Asn Tyr Thr Asp Ala Asp Gln Pro Xaa Gln Gln Ala Xaa Asp Thr  
 210 215 220  
 5 Ala Gly Thr Gln Ala Glu Ala Ile Thr Asn Ala Asn Gly Ser Xaa Ala  
 225 230 235 240  
 Asn Glu Thr Gln Val Gln Ala Ala Leu Asn Gln Leu Asn Gln Ala Lys  
 245 250 255  
 10 Asn Asp Leu Glu Trp Val Ile Ile Lys Leu Leu Lys Arg Lys Lys Gln  
 260 265 270  
 Gln Asn Val His  
 275

(2) INFORMATION FOR SEQ ID NO:517:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 234 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:517:

Val Asn Thr Thr Lys Ala Ala Leu His Gly Asp Val Lys Leu Gln Asn  
 1 5 10 15  
 35 Asp Lys Asp His Ala Lys Gln Thr Val Ser Gln Leu Ala His Leu Asn  
 20 25 30  
 Asn Ala Gln Lys His Met Glu Asp Thr Leu Ile Asp Ser Glu Thr Thr  
 35 40 45  
 40 Arg Thr Ala Val Lys Gln Asp Leu Thr Glu Val Gln Ala Leu Asp Gln  
 50 55 60  
 Leu Met Asp Ala Leu Gln Gln Ser Ile Ala Asp Lys Asp Ala Thr Arg  
 65 70 75 80  
 45 Ala Ser Ser Ala Tyr Val Asn Ala Glu Pro Asn Lys Lys Gln Ala Tyr  
 85 90 95  
 Asp Glu Ala Val Gln Asn Ala Glu Ser Ile Ile Ala Gly Leu Asn Asn  
 100 105 110  
 50 Pro Thr Ile Asn Lys Gly Asn Val Ser Ser Ala Thr Gln Ala Val Ile  
 115 120 125  
 Ser Ser Lys Asn Ala Leu Asp Gly Val Glu Arg Leu Ala Gln Asp Lys  
 130 135 140  
 55 Gln Thr Ala Gly Asn Ser Leu Asn His Leu Asp Gln Leu Thr Pro Ala  
 145 150 155 160

EP 0 841 394 A2

... Gln Gln Gln Ala Leu Glu Asn Gln Ile Asn Asn Ala Thr Thr Arg Asp  
 165 170 175  
 5 Lys Val Ala Glu Ile Ile Ala Gln Ala Gln Ala Leu Asn Glu Ala Met  
 180 185 190  
 Lys Ala Leu Lys Xaa Ser Ile Lys Asp Gln Pro Gln Thr Glu Ala Ser  
 195 200 205  
 10 Ser Lys Phe Ile Asn Glu Asp Gln Ala Gln Lys Asp Ala Tyr Thr Gln  
 210 215 220  
 Ala Val Gln His Ala Arg Arg Cys Leu Xaa  
 225 230

(2) INFORMATION FOR SEQ ID NO:518:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 55 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:518:

Val Ala Glu Ile Ile Ala Gln Ala Gln Ala Leu Asn Glu Ala Met Lys  
 1 5 10 15  
 Ala Leu Lys Glu Ser Ile Lys Asp Gln Pro Gln Thr Glu Ala Ser Ser  
 20 25 30  
 Lys Phe Ile Asn Glu Asp Gln Ala Gln Lys Asp Ala Tyr Thr Gln Ala  
 35 40 45  
 Val Gln His Glu Arg Lys Ile  
 50 55

(2) INFORMATION FOR SEQ ID NO:519:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 229 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:519:

EP 0 841 394 A2

```

... Val Val Val Leu Arg Met Ala Ser Asn Met Pro Ser Leu Xaa Lys Val ...
  1           5           10           15
5   Glu Asn Val Glu Met Ile Ala Phe Cys Asp Val Asp Ile Ser Lys Ala
    20           25           30
    Ala Ser Ala Ala Glu Ala Tyr Gly Thr Asp Asn Ala Lys Val Tyr Asp
      35           40           45
10  Asp Tyr Lys Ala Leu Leu Lys Asp Asp Thr Ile Asp Val Ile His Val
    50           55           60
    Cys Thr Pro Asn Asp Ser His Cys Glu Ile Thr Val Ala Gly Leu His
    65           70           75           80
15  Ala Gly Lys His Val Met Cys Glu Lys Pro Met Ala Lys Thr Thr Ala
    85           90           95
    Glu Ala Gln Lys Met Ile Asp Thr Ala Lys Ser Thr Gly Lys Lys Leu
    100          105          110
20  Thr Ile Gly Tyr Gln Asn Arg Phe Arg Pro Asp Ser Gln Phe Leu His
    115          120          125
    Gln Ala Ala Gln Arg Gly Asp Leu Gly Asp Ile Tyr Phe Gly Lys Ala
    130          135          140
25  His Ala Ile Arg Arg Arg Ala Val Pro Thr Trp Gly Val Phe Leu Asn
    145          150          155          160
    Glu Glu Ala Gln Gly Gly Gly Pro Leu Ile Asp Ile Gly Thr His Ala
    165          170          175
30  Leu Asp Leu Thr Leu Trp Met Met Asp Asn Tyr Glu Pro Glu Ser Val
    180          185          190
    Met Gly Ser Thr Phe His Lys Leu Asn Lys Gln His Asp Ala Pro Asn
    195          200          205
35  Ala Trp Gly Ser Trp Asn Pro Asp Glu Leu Thr Val Glu Asp Ser Ala
    210          215          220
    Phe Gly Phe Ile Lys
    225

```

(2) INFORMATION FOR SEQ ID NO:520:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 101 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:520:

EP 0 841 394 A2

... Val Asp Glu Ala Lys Cys Ser Leu Leu Gly Thr Lys Ala Gly Ala Asp ...  
 1 5 10 15  
 5 Met Lys Asp Val Leu Arg Ile His Gly Glu Asp Met Gly Thr Leu Tyr  
 20 25 30  
 Thr Lys His Val Glu Xaa Glu Asn Lys Gly Val Asp Phe Tyr Glu Gly  
 35 40 45  
 10 Asn Glu Val Asp Glu Ala Glu Glu Glu Xaa Lys Ala Trp Ile Asp Ala  
 50 55 60  
 Val Val Asn Asp Thr Glu Pro Val Val Lys Pro Glu Gln Ala Met Val  
 65 70 75 80  
 15 Val Thr Lys Ile Leu Glu Ala Ile Tyr Gln Ser Ala Lys Ser Gly Lys  
 85 90 95  
 Ala Ile Tyr Phe Glu  
 100

(2) INFORMATION FOR SEQ ID NO:521:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 61 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:521:

Val Asp Pro Pro Gly Cys Arg Asn Ser Thr Arg Thr Arg Ala Ala Phe  
 1 5 10 15  
 Thr Val Ala Ser Ile Asp Leu Gly Ala His Pro Glu Phe Leu Gly Lys  
 20 25 30  
 Asn Asp Ile Gln Leu Xaa Lys Lys Glu Ser Val Glu Asp Thr Xaa Lys  
 35 40 45  
 Val Leu Gly Arg Met Phe Asp Gly Ile Glu Phe Lys Leu  
 50 55 60

(2) INFORMATION FOR SEQ ID NO:522:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 123 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:522:

Val Ile Pro Asn Glu Phe Leu Lys His Ser Gly Lys Val His Ala Gln  
 1 5 10 15  
 Ala Phe Phe Thr Gln Asn Gly Ser Asn Asn Val Val Val Glu Arg Gln  
 20 25 30  
 Phe Ser Phe Asn Ile Glu Asn Asp Leu Val Ser Gly Xaa Asp Gly Ile  
 35 40 45  
 Thr Lys Leu Val Tyr Ile Lys Ser Ile Gln Asp Thr Ile Glu Ala Val  
 50 55 60  
 Gly Lys Asp Phe Asn Gln Leu Lys Gln Asn Met Ala Asp Thr Gln Thr  
 65 70 75 80  
 Leu Ile Ala Lys Val Asn Asp Ser Ala Thr Lys Gly Ile Gln Gln Ile  
 85 90 95  
 Glu Ile Lys Gln Asn Glu Ala Ile Gln Ala Ile Thr Ala Thr Gln Thr  
 100 105 110  
 Ser Ala Thr Gln Ala Val Thr Ala Glu Val Gly  
 115 120

(2) INFORMATION FOR SEQ ID NO:523:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 229 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:523:

Val Met Asn Ile Ile Asn Leu Glu Lys Pro Lys Gly Val Val Val Gln  
 1 5 10 15  
 Phe Gly Gly Gln Thr Ala Ile Asn Leu Ala Asp Lys Leu Ala Lys His  
 20 25 30  
 Gly Val Lys Ile Leu Gly Thr Ser Leu Glu Asn Leu Asn Arg Ala Glu  
 35 40 45  
 Asp Arg Lys Glu Phe Glu Ala Leu Leu Arg Lys Ile Asn Val Pro Gln  
 50 55 60  
 Pro Gln Gly Lys Ser Ala Thr Ser Pro Glu Glu Ala Leu Ala Asn Ala  
 65 70 75 80

EP 0 841 394 A2

Ala Glu Ile Gly Tyr Pro Val Val Val Arg Pro Ser Tyr Val Leu Gly  
 85 90 95  
 5 Gly Arg Ala Met Glu Ile Val Asp Asn Asp Lys Glu Leu Glu Asn Tyr  
 100 105 110  
 Met Thr Gln Ala Val Lys Ala Ser Pro Glu His Pro Val Leu Val Asp  
 115 120 125  
 10 Arg Tyr Leu Thr Gly Lys Glu Ile Glu Val Asp Ala Ile Cys Asp Gly  
 130 135 140  
 Glu Thr Val Ile Ile Pro Gly Ile Met Glu His Ile Glu Arg Ala Gly  
 145 150 155 160  
 15 Val His Ser Gly Asp Ser Ile Ala Val Tyr Pro Pro Gln Thr Leu Thr  
 165 170 175  
 Glu Asp Glu Leu Ala Thr Leu Glu Asp Tyr Thr Ile Lys Leu Ala Lys  
 180 185 190  
 20 Gly Leu Asn Ile Ile Gly Leu Ile Asn Ile Gln Phe Val Ile Ala His  
 195 200 205  
 Asp Gly Val Tyr Cys Phe Arg Ser Lys Pro Thr Val Leu Val Glu Arg  
 210 215 220  
 25 Xaa His Ser Arg Ala  
 225

(2) INFORMATION FOR SEQ ID NO:524:

30

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 90 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: Protein

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:524:

Val Ile Phe Met Thr Asn Asn Lys Val Ala Leu Val Thr Gly Gly Ala  
 45 1 5 10 15  
 Gln Gly Ile Gly Phe Lys Ile Ala Glu Arg Leu Val Glu Asp Gly Phe  
 20 25 30  
 Lys Val Ala Val Val Asp Phe Asn Glu Glu Gly Ala Lys Ala Ala Ala  
 50 35 40 45  
 Leu Lys Leu Ser Ser Asp Gly Thr Lys Ala Ile Ala Ile Lys Ala Asp  
 50 55 60  
 55 Val Ser Asn Arg Asp Asp Val Phe Asn Ala Val Arg Gln Ala Thr Pro  
 65 70 75 80



EP 0 841 394 A2

... Ser Ala Ile Trp Arg Phe Pro Cys His Gly...  
85 90

(2) INFORMATION FOR SEQ ID NO:525:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 191 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:525:

Val	Arg	Ile	His	Tyr	Val	Asp	Ile	Ile	Met	Phe	Lys	Asp	Val	Val	Ile
1			5					10					15		
Trp	Glu	Val	Cys	Leu	Ile	Arg	Leu	Leu	Val	Ile	Lys	Val	Leu	Lys	Arg
			20				25					30			
Met	Ile	Phe	Ala	Arg	Ala	Lys	Phe	Glu	Lys	Glu	Tyr	Gly	Val	Lys	Leu
		35				40					45				
Asn	Pro	Lys	Ala	Gly	Lys	Asp	Asn	His	Glu	Met	Val	Glu	Gly	Ile	His
	50				55				60						
Asp	Gly	Glu	Val	His	Ser	Leu	Tyr	Leu	Tyr	Gly	Glu	Asp	Thr	Gly	Ile
65				70				75					80		
Val	Asp	Ser	Asn	Ile	Asn	Phe	Val	Gln	Ala	Ala	Phe	Glu	Lys	Leu	Asp
			85				90						95		
Phe	Met	Val	Val	Gln	Asp	Glu	Phe	Leu	Thr	Phe	Thr	Ala	Thr	Tyr	Ala
			100				105					110			
Asp	Val	Val	Leu	Pro	Ala	Ser	Pro	Ser	Leu	Glu	Lys	Asp	Gly	Thr	Phe
		115				120					125				
Thr	Asn	Thr	Glu	Arg	Arg	Ile	Gln	Arg	Leu	Tyr	Gln	Ala	Leu	Glu	Pro
	130				135				140						
Leu	Gly	Asp	Ser	Lys	Pro	Asp	Trp	Lys	Ile	Phe	Gln	Ala	Ile	Ala	Asn
145				150				155					160		
Arg	Leu	Gly	Ile	Trp	Ile	Gly	Ile	Thr	Ser	Ile	Leu	Val	Lys	Leu	Trp
			165				170					175			
Ile	Glu	Gly	Arg	Thr	Leu	Asn	Thr	Ser	Ile	Cys	Trp	Gly	Lys	Leu	
		180				185						190			

(2) INFORMATION FOR SEQ ID NO:526:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 123 amino acids

EP 0 841 394 A2

(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:526:

```

Val Pro Ile Arg His Glu Arg Xaa Asp Lys Cys Glu Xaa Ile Met Tyr
 1             5             10             15
Ala Asp Asn Met Thr Asp Xaa Met Lys Tyr Ala Ile Asp Glu Thr Gln
15             20             25             30
Arg Arg Arg Glu Ile Gln Met Lys His Asn Glu Lys His Gly Ile Thr
35             40             45
Pro Lys Thr Ile Asn Lys Xaa Val His Asp Leu Ile Ser Ala Thr Val
20             50             55             60
Glu Asn Asp Glu Asn Xaa Asp Xaa Ala Gln Thr Val Ile Pro Lys Lys
65             70             75             80
Met Pro Lys Arg Ala Arg Xaa Lys Thr Ile Asp Asn Ile Glu Lys Glu
25             85             90             95
Met Lys Gln Ala Ala Lys Asp Leu Asp Phe Glu Lys Ala Xaa Glu Leu
100            105            110
Arg Asp Met Leu Phe Glu Leu Lys Ala Xaa Gly
30            115            120

```

(2) INFORMATION FOR SEQ ID NO:527:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 46 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:527:

```

Val Xaa Ala Asn Glu Ile Xaa Lys Xaa Lys Ile Asp Ala Asn Lys Asp
50 1             5             10             15
Val Asp Lys Gln Val Gln Ala Leu Ile Asp Glu Ile Asp Arg Asn Pro
20             25             30
Asn Leu Thr Asp Lys Glu Lys Gln Ala Leu Lys Arg Ser Tyr
55 35             40             45

```

EP 0 841 394 A2

(2) INFORMATION FOR SEQ ID NO:528:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 154 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:528:

```

Val Ser Glu Phe Asn Glu Ile Ile Ile Ala Thr Pro Ala Gln Trp Ile
 1             5             10             15
Ser His Thr Gln Asp Ile Leu Lys Lys Tyr Asn Ile Thr Asp Gln Arg
20             20             25             30
Val Lys Val Val Ala Gly Gly Xaa Asp Arg Asn Glu Thr Ile Met Asn
35             40             45
Ile Ile Asp His Ile Arg Asn Val Asn Gly Ile Asn Asn Asp Asp Val
25             50             55             60
Ile Val Thr His Asp Ala Val Arg Pro Phe Leu Thr Gln Arg Ile Ile
65             70             75             80
Lys Glu Asn Ile Glu Val Ala Xaa Xaa Tyr Gly Ala Val Asp Thr Val
30             85             90             95
Ile Glu Ala Ile Asp Thr Ile Val Met Ser Lys Asp Lys Gln Asn Ile
100            105            110
His Ser Ile Pro Val Arg Asn Glu Met Tyr Gln Gly Gln Thr Pro Gln
35            115            120            125
Ser Phe Asn Ile Lys Leu Leu Gln Asp Ser Tyr Arg Ala Leu Ser Ser
130            135            140
Glu Pro Xaa Arg Asn Leu Ile Arg Cys Met
40            145            150

```

(2) INFORMATION FOR SEQ ID NO:529:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 124 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:529:

EP 0 841 394 A2

Val Leu Phe Thr Ser Asp Ser Arg Pro Thr Met Ser Gly Asn Phe Glu  
 1 5 10 15  
 Tyr Ile Tyr Asn Glu Met Leu Arg Gln Asn Leu Asp Lys Lys Tyr Asp  
 5 20 25 30  
 Ile His Thr Val Phe Lys Ala Asn Ile Thr Asp Arg Arg Gly Ile Ile  
 35 40 45  
 Asp Lys Phe Arg Leu Pro Tyr Leu Leu Gly Lys Ala Asp Tyr Ile Phe  
 10 50 55 60  
 Val Asp Asp Phe His Pro Leu Ile Tyr Thr Val Arg Phe Arg Arg Ser  
 65 70 75 80  
 Gln Glu Val Ile Gln Val Trp His Ala Val Gly Ala Phe Lys Thr Val  
 15 85 90 95  
 Gly Phe Ser Arg Thr Gly Lys Lys Gly Gly Pro Phe Ile Asp Ser Leu  
 100 105 110  
 Asn His Arg Ser Ser Cys Gln Ser Leu Cys Ile Ile  
 20 115 120

(2) INFORMATION FOR SEQ ID NO:530:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 48 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:530:

Val Xaa His Ala Val Gly Ala Phe Lys Thr Val Gly Phe Ser Arg Thr  
 1 5 10 15  
 Gly Xaa Lys Gly Gly Pro Phe Ile Asp Ser Leu Lys Pro Ile Val Val  
 20 25 30  
 Thr Pro Lys Ala Tyr Val Ser Ser Glu Thr Arg Tyr Ser Ile Leu Cys  
 45 35 40 45

(2) INFORMATION FOR SEQ ID NO:531:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

EP 0 841 394 A2

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:531:

Val Tyr Phe Asp Leu Gly Lys Thr Ser Gly Ser Gly Thr Asn Ala Asn  
 1 5 10 15  
 Lys Val Thr Ile Thr Lys Ile Met Gly Trp Lys  
 20 25

(2) INFORMATION FOR SEQ ID NO:532:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 128 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:532:

Val Glu Ile Met Lys Ile Thr Val Asn Asp Lys Asn Glu Val Ile Gly  
 1 5 10 15  
 Tyr Val Asn Thr Gly Gly Leu Arg Asn Ser Leu Asp Val Asp Asp Asn  
 20 25 30  
 Asn Val Pro Ile Lys Phe Lys Glu Glu Phe Glu Pro Arg Lys Phe Val  
 35 40 45  
 Phe Thr Asn Gly Glu Ile Lys Tyr Asn Ser Asn Phe Glu Lys Glu Asp  
 50 55 60  
 Val Pro Asn Ala Ser Ser Gln Gln Ser Glu Ser Asp Leu Ser Asp Glu  
 65 70 75 80  
 Glu Leu Arg Gly Met Val Ala Ser Met Gln Met Gln Val Ala Gln Val  
 85 90 95  
 Asn Val Leu Thr Met Glu Leu Ala Gln Gln Asn Ala Met Leu Thr Gln  
 100 105 110  
 Gln Leu Thr Glu Leu Lys Thr Asn Lys Thr Ser Thr Glu Gly Asp Val  
 115 120 125

(2) INFORMATION FOR SEQ ID NO:533:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 377 amino acids
- (B) TYPE: amino acid

EP 0 841 394 A2

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:533:

10

15

20

25

30

35

40

45

50

55

```

Val Arg Gln Leu Ala Gln Ala Lys Lys Lys Ser Thr Ala Lys Lys Lys
 1           5           10           15
Thr Ala Ser Lys Lys Arg Thr Asn Ser Arg Lys Lys Lys Asn Asp Asn
           20           25           30
Pro Ile Arg Tyr Val Ile Ala Ile Leu Val Val Val Leu Met Val Leu
           35           40           45
Gly Val Phe Gln Leu Gly Ile Ile Gly Arg Leu Ile Asp Ser Phe Phe
           50           55           60
Asn Tyr Leu Phe Gly Tyr Ser Arg Tyr Leu Thr Tyr Ile Leu Val Leu
           65           70           75           80
Leu Ala Thr Gly Phe Ile Thr Tyr Ser Lys Arg Ile Pro Lys Thr Arg
           85           90           95
Arg Thr Ala Gly Ser Ile Val Leu Gln Ile Ala Leu Leu Phe Val Ser
           100          105          110
Gln Leu Val Phe His Phe Asn Ser Gly Ile Lys Ala Glu Arg Glu Pro
           115          120          125
Val Leu Ser Tyr Val Tyr Gln Ser Tyr Gln His Ser His Phe Pro Asn
           130          135          140
Phe Gly Gly Gly Val Leu Gly Phe Tyr Leu Leu Glu Leu Ser Val Pro
           145          150          155          160
Leu Ile Ser Leu Phe Gly Val Cys Ile Ile Thr Ile Leu Leu Leu Cys
           165          170          175
Ser Ser Val Ile Leu Leu Thr Asn His Gln His Arg Asp Val Ala Lys
           180          185          190
Val Ala Leu Glu Asn Ile Lys Ala Trp Phe Gly Ser Phe Asn Glu Lys
           195          200          205
Met Ser Glu Arg Asn Gln Glu Lys Gln Leu Lys Arg Glu Glu Lys Ala
           210          215          220
Arg Leu Lys Glu Glu Gln Lys Ala Arg Gln Asn Glu Gln Pro Gln Ile
           225          230          235          240
Lys Asp Val Ser Asp Phe Thr Glu Val Pro Gln Glu Arg Asp Ile Pro
           245          250          255
Ile Tyr Gly His Thr Glu Asn Glu Ser Lys Ser Gln Cys Gln Pro Ser
           260          265          270
Arg Lys Lys Arg Val Phe Asp Ala Glu Asn Ser Ser Asn Asn Ile Val
           275          280          285

```

# EP 0 841 394 A2

Asn His Gln Ala Asp Gln Gln Glu Gln Leu Thr Glu Gln Thr His Asn  
 290 295 300  
 Ser Val Glu Ser Glu Asn Thr Ile Glu Glu Ala Gly Glu Val Thr Asn  
 5 305 310 315 320  
 Val Ser Tyr Val Val Pro Pro Leu Thr Leu Leu Asn Gln Pro Ala Lys  
 325 330 335  
 Gln Lys Ala Thr Ser Lys Ala Glu Val Gln Arg Lys Gly Gln Val Pro  
 10 340 345 350  
 Lys Arg Tyr Xaa Lys Arg Phe Trp Gly Xaa Ile Xaa Lys Xaa Thr Gln  
 355 360 365  
 Xaa Lys Leu Ser Xaa Ser Asn Ser Ile  
 15 370 375

## (2) INFORMATION FOR SEQ ID NO:534:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 186 amino acids  
 (B) TYPE: amino acid  
 25 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:534:

Val Met Thr Glu Glu Xaa Cys Leu Leu Lys Arg Val Phe Met Arg Met  
 35 1 5 10 15  
 Lys Arg Phe Leu Thr Ile Val Gln Ile Leu Leu Val Val Ile Ile Ile  
 20 25 30  
 Ile Phe Gly Tyr Lys Ile Val Gln Thr Tyr Ile Glu Asp Lys Gln Glu  
 40 35 40 45  
 Arg Ala Xaa Tyr Glu Lys Leu Gln Xaa Lys Phe Xaa Met Leu Met Xaa  
 50 55 60  
 Lys His Gln Glu His Val Arg Pro Gln Phe Glu Ser Leu Glu Lys Ile  
 45 65 70 75 80  
 Asn Lys Asp Ile Val Gly Trp Ile Lys Leu Ser Gly Thr Ser Leu Xaa  
 85 90 95  
 Tyr Pro Val Leu Gln Gly Lys Thr Asn His Asp Tyr Leu Asn Leu Asp  
 50 100 105 110  
 Phe Glu Arg Glu His Arg Arg Lys Gly Ser Ile Phe Met Asp Phe Arg  
 115 120 125  
 Asn Glu Leu Xaa Asn Leu Asn His Asn Thr Ile Leu Tyr Gly His His  
 55 130 135 140

Val Gly Asp Asn Thr Met Phe Asp Val Leu Glu Asp Tyr Leu Lys Gln  
 145 150 155 160  
 5 Ser Phe Tyr Glu Lys His Lys Ile Ile Gly Phe Asp Asn Lys Tyr Gly  
 165 170 175  
 Lys Tyr Gln Leu Gln Val Phe Ser Ala Tyr  
 180 185  
 10

# Claims

1. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:
  - (a) a polynucleotide having at least a 70% identity to a polynucleotide encoding a polypeptide comprising an amino acid sequence of the Sequence Listing;
  - (b) a polynucleotide having at least a 70% identity to a polynucleotide encoding a mature polypeptide expressed by the gene contained in the *S. aureus* of the deposited strain that was sequenced to obtain a polynucleotide sequence of the Sequence Listing;
  - (c) a polynucleotide encoding a polypeptide comprising an amino acid sequence which is at least 70% identical to an amino acid sequence of the Sequence Listing;
  - (d) a polynucleotide which is complementary to the polynucleotide of (a), (b) or (c); and
  - (e) a polynucleotide comprising at least 15 sequential bases of the polynucleotide of (a), (b), (c) or (d).
2. The polynucleotide of Claim 1 wherein the polynucleotide is DNA.
3. The polynucleotide of Claim 1 wherein the polynucleotide is RNA.
4. The polynucleotide of Claim 2 comprising the nucleic acid sequence selected from the group consisting of the nucleic acid sequences set forth in the Sequence Listing.
5. The polynucleotide of Claim 2 which encodes a polypeptide comprising an amino acid sequence selected from the group consisting of the amino acid sequences set forth in the Sequence Listing.
6. A vector comprising the polynucleotide of Claim 1.
7. A host cell comprising the vector of Claim 6.
8. A process for producing a polypeptide comprising: expressing from the host cell of Claim 7 a polypeptide encoded by said DNA.
9. A process for producing a polypeptide or fragment comprising culturing a host of claim 7 under conditions sufficient for the production of said polypeptide or fragment.
10. A polypeptide comprising an amino acid sequence which is at least 70% identical to an amino acid sequence selected from the group consisting of the amino acid sequences set forth in the Sequence Listing.
11. A polypeptide comprising an amino acid sequence selected from the group consisting of the amino acid sequences set forth in the Sequence Listing.
12. An antibody against the polypeptide of claim 10.
13. An antagonist or agonist of the activity or expression of the polypeptide of claim 10.
14. A method for the treatment or prevention of disease of an individual comprising: administering to the individual a therapeutically effective amount of the polypeptide of claim 10.



15. A method for the treatment of an individual having need to inhibit a bacterial polypeptide comprising: administering to the individual a therapeutically effective amount of the antagonist of Claim 13.

16. A process for diagnosing a disease related to expression or activity of the polypeptide of claim 10 in an individual comprising:

- (a) determining a nucleic acid sequence encoding said polypeptide, and/or
- (b) analyzing for the presence or amount of said polypeptide in a sample derived from the individual.

17. A method for identifying compounds which interact with and inhibit or activate an activity of the polypeptide of claim 10 comprising:

contacting a composition comprising the polypeptide with the compound to be screened under conditions to permit interaction between the compound and the polypeptide to assess the interaction of a compound, such interaction being associated with a second component capable of providing a detectable signal in response to the interaction of the polypeptide with the compound;  
and determining whether the compound interacts with and activates or inhibits an activity of the polypeptide by detecting the presence or absence of a signal generated from the interaction of the compound with the polypeptide.

18. A method for inducing an immunological response in a mammal which comprises inoculating the mammal with the polypeptide of claim 10, or a fragment or variant thereof, adequate to produce antibody and/or T cell immune response to protect said animal from disease.

19. A method of inducing immunological response in a mammal which comprises delivering a nucleic acid vector to direct expression of a polypeptide of claim 10, or fragment or a variant thereof, for expressing said polypeptide, or a fragment or a variant thereof *in vivo* in order to induce an immunological response to produce antibody and/or T cell immune response to protect said animal from disease.

20. A polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO: 1, 52, 95, 184, 127, 128, 153, 212, 215 and 227.

21. A polypeptide comprising a polypeptide encoded by the polynucleotide of claim 20.

22. The isolated polynucleotide of claim 1 wherein said nucleotide is selected from the group consisting of:

- (a) a polynucleotide having at least a 90% identity to a polynucleotide encoding a polypeptide comprising the amino acid sequence of the Sequence Listing;
- (b) a polynucleotide having at least a 90% identity to a polynucleotide encoding the same mature polypeptide expressed by the gene contained in the *S. aureus* of the deposited strain that was sequenced to obtain a polynucleotide sequence of the Sequence Listing;
- (c) a polynucleotide encoding a polypeptide comprising an amino acid sequence which is at least 90% identical to the amino acid sequence of the Sequence Listing;
- (d) a polynucleotide which is complementary to the polynucleotide of (a), (b) or (c); and
- (e) a polynucleotide comprising at least 15 sequential bases of the polynucleotide of (a), (b), (c) or (d).

23. The isolated polynucleotide of claim 1 selected from the group consisting of:

- (a) a polynucleotide having at least a 95% identity to a polynucleotide encoding a polypeptide comprising the amino acid sequence of the Sequence Listing;
- (b) a polynucleotide having at least a 95% identity to a polynucleotide encoding the same mature polypeptide expressed by the gene contained in the *S. aureus* of the deposited strain that was sequenced to obtain a polynucleotide sequence of the Sequence Listing;
- (c) a polynucleotide encoding a polypeptide comprising an amino acid sequence which is at least 95% identical to the amino acid sequence of the Sequence Listing;
- (d) a polynucleotide which is complementary to the polynucleotide of (a), (b) or (c); and
- (e) a polynucleotide comprising at least 15 sequential bases of the polynucleotide of (a), (b), (c) or (d).

24. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

- (a) a polynucleotide having at least a 50% identity to a polynucleotide encoding a polypeptide comprising the amino acid sequence of the Sequence Listing and obtained from a prokaryotic species other than *S. aureus*;  
(b) a polynucleotide encoding a polypeptide-comprising an amino acid sequence which is at least 50% identical to the amino acid sequence of the Sequence Listing and obtained from a prokaryotic species other than *S. aureus*; and  
(c) a polynucleotide which is complementary to the polynucleotide of (a) or (b).

25. An isolated Streptococcal polypeptide having one of the amino acid sequences given in Table 1.

26. An isolated nucleic acid encoding one of the amino acid sequences of Claim 1 and nucleic acid sequences capable of hybridizing therewith under stringent conditions.

27. Recombinant vectors comprising the nucleic acid sequences of Claim 26 and host cells transformed or transfected therewith.

28. A method of identifying an antimicrobial compound comprising contacting candidate compounds with a polypeptide of Claim 1 and selecting those compounds capable of inhibiting the bioactivity of said polypeptide.

29. Antimicrobial compounds identified by the method of Claim 28.

(19)



Europäisches Patentamt  
European Patent Office  
Office européen des brevets

(11) **EP 0 841 394 A3**

(12)

**EUROPEAN PATENT APPLICATION**

(88) Date of publication A3:  
02.02.2000 Bulletin 2000/05

(43) Date of publication A2:  
13.05.1998 Bulletin 1998/20

(51) Int. **C12N 15/31**, C07K 14/31,  
Cl.7: C07K 16/12, C12N 1/21,  
A61K 38/16, G01N 33/50,  
C12Q 1/68

(21) Application number: **97307485.9**

(22) Date of filing: **24.09.1997**

(84) Designated Contracting States:  
**AT BE CH DE DK ES FI FR GB GR IE IT LI  
LU MC NL PT SE**

(30) Priority: **24.09.1996 US 27032 P**

(71) Applicants:  
, **SMITHKLINE BEECHAM CORPORATION**  
Philadelphia Pennsylvania 19103 (US)  
, **SMITHKLINE BEECHAM PLC**  
Brentford, Middlesex TW8 9EP (GB)

(72) Inventors:  
, **Black, Michael T., SmithKline Beecham**  
**Pharma.**  
Collegeville, Pennsylvania 19426-  
0989 (US)  
, **Hodgson, John Edward, SmithKline**  
**Beecham Pharma.**  
Collegeville, Pennsylvania 19426-  
0989 (US)  
, **Knowles, David J. C., SmithKline**  
**Beecham Pharma.**  
Harlow, Essex CM19 5AW (GB)  
, **Reichard, Raymond W., SmithKline**  
**Beecham Pharma.**  
King of Prussia, Pennsylvania  
19406 (US)

, **Nicholas, Richard O., SmithKline**  
**Beecham Pharma.**  
Collegeville, Pennsylvania 19426-  
0989 (US)  
, **Burnham, Martin K. R., SmithKline**  
**Beecham Pharma.**  
Collegeville, Pennsylvania 19426-  
0989 (US)  
, **Pratt, Julie M., SmithKline Beecham**  
**p.l.c.**  
Harlow, Essex CM19 5AW (GB)  
, **Rosenberg, Martin, SmithKline**  
**Beecham Pharma.**  
King of Prussia, Pennsylvania  
19406 (US)  
, **Ward, Judith M., SmithKline Beecham**  
**Pharma.**  
Harlow, Essex CM19 5AW (GB)  
, **Lonetto, Michael A., SmithKline**  
**Beecham Pharma.**  
Collegeville, Pennsylvania 19426-  
0989 (US)

(74) Representative:  
**Mallalieu, Catherine Louise et al**  
**D. Young & Co.,**  
21 New Fetter Lane  
London EC4A 1DA (GB)

(54) **Staphylococcus aureus polynucleotides, polypeptides and their uses**

(57) The invention provides novel  
polypeptides and polynucleotides encoding  
such polypeptides and methods for producing

such polypeptides by recombinant  
techniques. Also provided are methods for  
utilizing such polypeptides to screen for  
antibacterial compounds.

**EP 0 841 394 A3**

EP 0 841 394 A3



European Patent  
Office

**PARTIAL EUROPEAN SEARCH REPORT**

Application Number

which under Rule 45 of the European Patent Convention EP 97 30 7485  
shall be considered, for the purposes of subsequent  
proceedings, as the European search report

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
X	WO 94 01583 A (FUSO PHARMACEUTICAL IND ;EDA SOJI (JP); MATSUHISA AKIO (JP); OHNO) 20 January 1994 (1994-01-20) * abstract * * nt 1-396 of seq. ID 1 of the present application are complementary to nt 9535-9921 of probe SA-24 (W09401583, Seq. ID 2) *	1-29	C12N15/31 C07K14/31 C07K16/12 C12N1/21 A61K38/16 G01N33/50 C12Q1/68
P,X	EP 0 786 519 A (HUMAN GENOME SCIENCES INC) 30 July 1997 (1997-07-30) * abstract * * compare nt 8-396 of seq. ID 1 of the present application with nt 4326-4711 of EP786519, seq. ID 505 * claims 1-29 *	1-29	
A	WO 94 06830 A (ALFA LAVAL AGRI INT ;BODEN MARIA K (SE); FLOCK JAN INGMAR (SE)) 31 March 1994 (1994-03-31) * abstract *	1-29	
			TECHNICAL FIELDS SEARCHED (Int.Cl.6)
			C07K C12N A61K G01N C12Q
<b>INCOMPLETE SEARCH</b>			
<p>The Search Division considers that the present application, or one or more of its claims, does/do not comply with the EPC to such an extent that a meaningful search into the state of the art cannot be carried out, or can only be carried out partially, for these claims.</p> <p>Claims searched completely :</p> <p>Claims searched incompletely :</p> <p>Claims not searched :</p> <p>Reason for the limitation of the search:</p> <p>see sheet C</p>			
Place of search		Date of completion of the search	Examiner
THE HAGUE		7 September 1999	Galli, I
CATEGORY OF CITED DOCUMENTS		<p>T : theory or principle underlying the invention</p> <p>E : earlier patent document, but published on, or after the filing date</p> <p>D : document cited in the application</p> <p>L : document cited for other reasons</p> <p>&amp; : member of the same patent family, corresponding document</p>	
<p>X : particularly relevant if taken alone</p> <p>Y : particularly relevant if combined with another document of the same category</p> <p>A : technological background</p> <p>O : non-written disclosure</p> <p>P : intermediate document</p>			

EPO FORM 1503 (02.82) (P04007)



European Patent  
Office

INCOMPLETE SEARCH  
SHEET C

Application Number  
EP 97 30 7485

Although claims 14,15,18,19 are directed to methods of treatment of the human/animal body (Article 52(4) EPC), the search has been carried out and based on the alleged effects of the compound/composition.

-----

Claim(s) searched incompletely:  
13,15,29

Reason for the limitation of the search:

The antagonists of claims 13 and 15 and the antimicrobial compounds of claim 19 are not sufficiently described to allow for a complete and meaningful search.



European Patent  
Office

Application Number

EP 97 30 7485

### CLAIMS INCURRING FEES

The present European patent application comprised at the time of filing more than ten claims.

☐ Only part of the claims have been paid within the prescribed time limit. The present European search report has been drawn up for the first ten claims and for those claims for which claims fees have been paid, namely claim(s):

☐ No claims fees have been paid within the prescribed time limit. The present European search report has been drawn up for the first ten claims.

### LACK OF UNITY OF INVENTION

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

see sheet B

☐ All further search fees have been paid within the fixed time limit. The present European search report has been drawn up for all claims.

☐ As all searchable claims could be searched without effort justifying an additional fee, the Search Division did not invite payment of any additional fee.

☐ Only part of the further search fees have been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the inventions in respect of which search fees have been paid, namely claims:

☒ None of the further search fees have been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the invention first mentioned in the claims, namely claims:

See additional sheet, Invention 1.



European Patent  
Office

LACK OF UNITY OF INVENTION  
SHEET B

Application Number  
EP 97 30 7485

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

1. Claims: 1-29

Invention 1:  
-----

(claims 1-29, partially)

An isolated polynucleotide comprising a sequence homologous or identical to the *S. aureus* sequence of Seq. ID 1.

Corresponding vectors, recombinant cell, polypeptides (Seq. ID 260) and fragments thereof, antibodies, antagonists, antimicrobials, probes, pharmaceutical and diagnostic uses.

2. Claims: 1-29

Inventions 2-259:  
-----

(claims 1-29, partially)

Idem as subject matter 1, but wherein

invention 2 is limited to seq. IDs 2 and 261

invention 3 is limited to seq. IDs 3 and 262

etc.

invention 259 is limited to seq. IDs 259 and 534.

(For the sake of conciseness, inventions 2-259 are not listed singly.)



# EP 0 841 394 A3

## ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO.

EP 97 30 7485

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.  
The members are as contained in the European Patent Office EDP file on  
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

07-09-1999

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9401583 A	20-01-1994	AU 684250 B	11-12-1997
		AU 4513593 A	31-01-1994
		EP 0652291 A	10-05-1995
		JP 2965543 B	18-10-1999
		JP 10304895 A	17-11-1998
		JP 2965544 B	18-10-1999
		JP 10304896 A	17-11-1998
		JP 10304897 A	17-11-1998
		JP 2798499 B	17-09-1998
		US 5807673 A	15-09-1998
		US 5763188 A	09-06-1998
		US 5770375 A	23-06-1998
		US 5853998 A	29-12-1998
		US 5798211 A	25-08-1998
EP 0786519 A	30-07-1997	CA 2194411 A	06-07-1997
		JP 9322781 A	16-12-1997
WO 9406830 A	31-03-1994	AU 4837193 A	12-04-1994
		CA 2124210 A	31-03-1994
		EP 0621875 A	02-11-1994

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82